Уральский
федеральный
университет
имени первого Президента
России Б. Н. Ельцина
Институт естественных наук
и математики

А. Я. ОВСЯННИКОВ

ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ ТЕОРИИ МАТРИЦ

Учебное пособие

МИНИСТЕРСТВО НАУКИ И ВЫСШЕГО ОБРАЗОВАНИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ
УРАЛЬСКИЙ ФЕДЕРАЛЬНЫЙ УНИВЕРСИТЕТ
ИМЕНИ ПЕРВОГО ПРЕЗИДЕНТА РОССИИ Б. Н. ЕЛЬЦИНА

А. Я. Овсянников

ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ ТЕОРИИ МАТРИЦ

Учебное пособие

Рекомендовано
методическим советом Уральского федерального университета
в качестве учебного пособия для студентов вуза, обучающихся
по направлениям подготовки 01.03.01 «Математика»,
01.03.03 «Механика и математическое моделирование»,
02.03.01 «Математика и компьютерные науки»,
02.03.02 «Фундаментальная информатика
и информационные технологии»

Екатеринбург
Издательство Уральского университета
2020

УДК 512.643(075.8)
О345

Рецензенты:
кафедра высшей математики и методики обучения математике
Уральского государственного педагогического университета
(заведующий кафедрой доктор физико-математических наук,
доцент В. Ю. Бодряков);
Е. А. Перминов, доктор педагогических наук, кандидат физико-математических наук (Российский государственный
профессионально-педагогический университет)

Научный редактор
доктор физико-математических наук Б. М. Верников

Овсянников, А. Я.
О345 Дополнительные главы теории матриц : учеб. пособие / А. Я. Овсянников ; науч. ред. Б. М. Верников ; Министерство науки и высшего образования Российской Федерации, Уральский федеральный университет. – Екатеринбург : Изд-во Урал. ун-та, 2020. – 107 с. – 30 экз. – ISBN 978-5-7996-2947-2. – Текст : непосредственный.
ISBN 978-5-7996-2947-2

Учебное пособие содержит изложение некоторых важных разделов теории матриц, которые обычно недостаточно освещаются в стандартных курсах линейной алгебры.

Для студентов и преподавателей математических специальностей и направлений.

УДК 512.643(075.8)

ISBN 978-5-7996-2947-2

Оглавление

Предисловие 4
Введение 5
Глава 1. Определители, связанные с матрицей 7
§ 1.1. Формула Бине–Коши и ее приложения 7
§ 1.2. О решениях матричных уравнений 11
§ 1.3. Ассоциированная матрица 13
§ 1.4. Выражение миноров обратной матрицы через миноры исходной матрицы 15
§ 1.5. Псевдообратная матрица 16
Задачи для самостоятельного решения 25
Глава 2. Алгоритм Гаусса и его приложения 25
§ 2.1. Метод Гаусса 25
§ 2.2. Детерминантное тождество Сильвестра 27
§ 2.3. Разложение в произведение треугольных матриц 29
§ 2.4. Блочные матрицы 35
Задачи для самостоятельного решения 44
Глава 3. Многочленные матрицы 44
§ 3.1. Элементарные преобразования 44
§ 3.2. Делители миноров 47
§ 3.3. Элементарные делители 52
Задачи для самостоятельного решения 56
Глава 4. Подобие матриц 57
§ 4.1. Матричные многочлены 57
§ 4.2. Скалярная эквивалентность. Критерий подобия 59
§ 4.3. Нормальные формы 62
§ 4.4. Диагонализируемые матрицы 65
§ 4.5. Матрицы, перестановочные с данной матрицей 69
Задачи для самостоятельного решения 72

Глава 5. Характеристический и минимальный многочлены матрицы . . . 74
§ 5.1. Характеристический многочлен . . . 74
§ 5.2. Метод Д. К. Фаддеева . . . 77
§ 5.3. Минимальный многочлен . . . 79
Задачи для самостоятельного решения . . . 83
Глава 6. Прямое произведение матриц . . . 84
§ 6.1. Определение и свойства прямого произведения матриц . . . 84
§ 6.2. Составные матрицы . . . 87
Задачи для самостоятельного решения . . . 92
Глава 7. Функции от матриц . . . 92
§ 7.1. Определение функции от матрицы . . . 92
§ 7.2. Компоненты матрицы . . . 97
§ 7.3. Последовательности и ряды матриц . . . 101
Задачи для самостоятельного решения . . . 106

Предисловие

Учебное пособие посвящено изложению некоторых разделов теории матриц, которые обычно не освещаются в курсах линейной алгебры для студентов-бакалавров математических, физических и экономических специальностей. Однако эти разделы весьма интересны и важны для формирования у студентов полноценного представления о линейной алгебре и теории матриц. Пособие включает информацию о минорах матриц, приложения метода Гаусса, теорию многочленных матриц, критерий подобия матриц, понятия о прямом произведении матриц и функциях от матриц. Наряду с теоретическим материалом пособие содержит задачи для самостоятельного решения. Оно может служить основой односеместрового спецкурса для студентов 3-го или 4-го курса.

Введение

Линейная алгебра представляет собой фундаментальную математическую дисцилину для математиков, физиков и экономистов. Одной из основных ветвей линейной алгебры является теория матриц, которой посвящено большое количество монографий и учебников. Но в курсах линейной алгебры для студентов-бакалавров многие важные разделы теории матриц не могут быть изложены ввиду ограниченности объема этих курсов. В данном пособии представлены некоторые такие разделы. При его составлении были использованы монографии Р. Беллмана, Ф. Р. Гантмахера, П. Ланкастера, Р. Хорна и Ч. Джонсона, а также университетские учебники А. И. Мальцева и И. В. Проскурякова*.

В пособии рассматриваются матрицы над ассоциативно-коммутативными кольцами. При необходимости кольца заменяются полями или даже полями действительных или комплексных чисел. Матрицы записываются жирным шрифтом и обозначаются с помощью круглых скобок.

Определитель квадратной матрицы $\boldsymbol{A}$ обозначается через $|\boldsymbol{A}|$, а ее след – через $\mathsf{tr}(\boldsymbol{A})$. Ранг произвольной матрицы $\boldsymbol{A}$ обозначается через $\mathsf{r}(\boldsymbol{A})$.

Для матрицы $\boldsymbol{A} = (a_{ij})_{m\times n}$ через $\boldsymbol{A}_{i\bullet}$ обозначается ее i-я строка, а через $\boldsymbol{A}_{\bullet j}$ – j-й столбец: $\boldsymbol{A}_{i\bullet} = (a_{i1}a_{i2}\ldots a_{in})$,

---

** Беллман Р.* Основы теории матриц. М. : Наука, 1969; *Гантмахер Ф. Р.* Теория матриц. 4-е изд., доп. М. : Наука, 2010; *Ланкастер П.* Теория матриц. М. : Наука, 1978; *Мальцев А. И.* Основы линейной алгебры. М. : Наука, 1985; *Проскуряков И.В.* Сборник задач по линейной алгебре. М. : БИНОМ, 2005. (Классический университетский учебник); *Хорн Р., Джонсон Ч.* Матричный анализ. М. : Мир, 1989.

$\boldsymbol{A}_{\bullet j} = (a_{1j} a_{2j} \ldots a_{mj})^{\top}$. Элемент a_{ij} матрицы $\boldsymbol{A}$ будем иногда обозначать через $\boldsymbol{A}[i, j]$.

Диагональную матрицу порядка n с элементами $a_1, a_2, \ldots, a_n$ на главной диагонали будем обозначать через $[a_1, a_2, \ldots, a_n]$.

Нумерация утверждений (теорем, предложений, лемм и следствий) сплошная. Формулы нумеруются в пределах параграфа.

Каждая глава завершается набором примеров вычислительного характера для самостоятельного решения.

Глава 1

Определители, связанные с матрицей

§ 1.1. Формула Бине–Коши и ее приложения

Зафиксируем матрицу размеров $m \times n$:

$$\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{m1} & a_{m2} & \dots & a_{mn} \end{pmatrix} \tag{1}$$

над ассоциативно-коммутативным кольцом F. Пусть k – натуральное число, $1 \le i_1, \dots, i_k \le m$, $1 \le j_1, \dots, j_k \le n$. Для матрицы $\boldsymbol{A}$ из равенства (1) обозначим через

$$\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \dots & i_k \\ j_1 & j_2 & \dots & j_k \end{pmatrix} = \begin{vmatrix} a_{i_1 j_1} & a_{i_1 j_2} & \dots & a_{i_1 j_k} \\ a_{i_2 j_1} & a_{i_2 j_2} & \dots & a_{i_2 j_k} \\ \dots & \dots & \dots & \dots \\ a_{i_1 j_k} & a_{i_2 j_k} & \dots & a_{i_k j_k} \end{vmatrix}$$

определитель порядка k, составленный из элементов матрицы $\boldsymbol{A}$.

Здесь индексы $i_1, \dots, i_k$ или $j_1, \dots, j_k$ не обязательно идут в порядке возрастания и среди них могут встретиться одинаковые числа. В частности, при $k > m$ или $k > n$ всегда $\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \dots & i_k \\ j_1 & j_2 & \dots & j_k \end{pmatrix} = 0$.

Наблюдение 1. *При* $1 \le i_1 < i_2 < \ldots < i_k \le m$ *и* $1 \le j_1 < j_2 < \ldots < j_k \le n$ *определитель* $\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \dots & i_k \\ j_1 & j_2 & \dots & j_k \end{pmatrix}$ *является минором порядка* k *матрицы* $\boldsymbol{A}$.

В частности, для квадратной матрицы $\boldsymbol{A}$ порядка n

$$|\boldsymbol{A}| = \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & n \\ 1 & 2 & \dots & n \end{pmatrix}.$$

Через $\mathsf{ad}\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \dots & i_k \\ j_1 & j_2 & \dots & j_k \end{pmatrix}$ обозначим алгебраическое дополнение в квадратной матрице $\boldsymbol{A}$ порядка n к минору $\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \dots & i_k \\ j_1 & j_2 & \dots & j_k \end{pmatrix}$ $(k < n)$. Тогда имеет место формула

$$\mathsf{ad}\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \dots & i_k \\ j_1 & j_2 & \dots & j_k \end{pmatrix} = \\ = (-1)^{\sum_{t=1}^{k}(i_t+j_t)}\boldsymbol{A}\begin{pmatrix} i'_1 & i'_2 & \dots & i'_{n-k} \\ j'_1 & j'_2 & \dots & j'_{n-k} \end{pmatrix}, \qquad (2)$$

где $1 \le i_1 < i_2 < \ldots < i_k \le n$ вместе с $1 \le i'_1 < i'_2 < \ldots < i'_{n-k} \le n$ и $1 \le j_1 < j_2 < \ldots < j_k \le n$ вместе с $1 \le j'_1 < j'_2 < \ldots < j'_{n-k} \le n$ образуют полную систему индексов $\{1, 2, \ldots, n\}$.

Напомним, что через $F^{k\times n}$ обозначается кольцо матриц размеров $k \times n$ над кольцом F. В следующем утверждении доказывается важная *формула Бине–Коши*.

Теорема 1. *Пусть* $\boldsymbol{A} \in F^{m\times n}$, $\boldsymbol{B} \in F^{n\times m}$ *и* $\boldsymbol{C} = \boldsymbol{A}\cdot\boldsymbol{B}$. *Тогда*

$$|\boldsymbol{C}| = \sum_{1\le k_1<k_2<\ldots<k_m\le n} \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix}\boldsymbol{B}\begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}. \qquad (3)$$

Согласно формуле Бине–Коши, при $m < n$ определитель матрицы $\boldsymbol{C}$ равен сумме произведений всевозможных миноров максимального (m-го) порядка матрицы $\boldsymbol{A}$ на соответствующие миноры того же порядка матрицы $\boldsymbol{B}$; при $m > n$ в правой части формулы

Бине–Коши получается нуль. При $m = n$ формула Бине–Коши дает новое доказательство того, что определитель произведения двух квадратных матриц равен произведению их определителей.

Доказательство. Пусть $\boldsymbol{A} = (a_{ij})$, $\boldsymbol{B} = (b_{ij})$. Запишем по определению

$$|\boldsymbol{C}| = \begin{vmatrix} \sum_{t_1=1}^n a_{1t_1}b_{t_1 1} & \sum_{t_2=1}^n a_{1t_2}b_{t_2 2} & \dots & \sum_{t_m=1}^n a_{1t_m}b_{t_m m} \\ \sum_{t_1=1}^n a_{2t_1}b_{t_1 1} & \sum_{t_2=1}^n a_{2t_2}b_{t_2 2} & \dots & \sum_{t_m=1}^n a_{2t_m}b_{t_m m} \\ \dots & \dots & \dots & \dots \\ \sum_{t_1=1}^n a_{mt_1}b_{t_1 1} & \sum_{t_2=1}^n a_{mt_2}b_{t_2 2} & \dots & \sum_{t_m=1}^n a_{mt_m}b_{t_m m} \end{vmatrix}.$$

Раскладываем последний определитель в сумму по каждому столбцу и из каждого столбца выносим множитель $b_{t_j j}$:

$$|\boldsymbol{C}| = \sum_{t_1,t_2,\dots,t_m=1}^{n} \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & m \\ t_1 & t_2 & \dots & t_m \end{pmatrix} b_{t_1 1}b_{t_2 2}\dots b_{t_m m}. \qquad (4)$$

Если $m > n$, то при любых $t_1, t_2, \dots, t_m \in \{1, 2, \dots, n\}$ справедливо $\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & m \\ t_1 & t_2 & \dots & t_m \end{pmatrix} = 0$, и формула доказана.

Пусть $m \leq n$. Сумма в правой части (4) после отбрасывания нулевых слагаемых, возникающих, когда $t_i = t_j$ при $i \neq j$, превращается в сумму

$\sum_{1\leq k_1<k_2<\dots<k_m\leq n}(\sum_{\{t_1,t_2,\dots,t_m\}=\{k_1,k_2,\dots,k_m\}} \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & m \\ t_1 & t_2 & \dots & t_m \end{pmatrix}\cdot$
$\cdot b_{t_1 1}b_{t_2 2}\dots b_{t_m m})$.

Пусть $I(t_1, t_2, \dots, t_m)$ – число инверсий в перестановке $(t_1, t_2, \dots, t_m)$. Тогда при $\{t_1, t_2, \dots, t_m\} = \{k_1, k_2, \dots, k_m\}$ и $1 \leq k_1 < k_2 < \dots < k_m \leq n$ в силу свойств определителей $\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & m \\ t_1 & t_2 & \dots & t_m \end{pmatrix} = (-1)^{I(t_1,t_2,\dots,t_m)}\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix}$.

Итак, при фиксированных индексах $1 \leq k_1 < k_2 < \dots < k_m \leq n$ имеем

$$\sum_{\{t_1,t_2,\dots,t_m\}=\{k_1,k_2,\dots,k_m\}} \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & m \\ t_1 & t_2 & \dots & t_m \end{pmatrix} b_{t_1 1}b_{t_2 2}\dots b_{t_m m} =$$

$$= \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} \cdot$$
$$\cdot \sum_{\{t_1,t_2,\dots,t_m\}=\{k_1,t_2,\dots,k_m\}} (-1)^{I(t_1,t_2,\dots,t_m)} b_{t_1 1} b_{t_2 2} \dots b_{t_m m} =$$
$$= \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & m \\ k_1 & k_2 & \dots & k_m \end{pmatrix} \boldsymbol{B}\begin{pmatrix} k_1 & k_2 & \dots & k_m \\ 1 & 2 & \dots & m \end{pmatrix}.$$

Этим доказательство завершается.

Рассмотрим примеры.

1. Из матричного равенства

$$\begin{pmatrix} a_1c_1 + a_2c_2 + \dots + a_nc_n & a_1d_1 + a_2d_2 + \dots + a_nd_n \\ b_1c_1 + b_2c_2 + \dots + b_nc_n & b_1d_1 + b_2d_2 + \dots + b_nd_n \end{pmatrix} =$$
$$= \begin{pmatrix} a_1 & a_2 & \dots & a_n \\ b_1 & b_2 & \dots & b_n \end{pmatrix} \cdot \begin{pmatrix} c_1 & d_1 \\ c_2 & d_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ c_n & d_n \end{pmatrix}$$

получается *тождество Коши*

$$\begin{vmatrix} a_1c_1 + a_2c_2 + \dots + a_nc_n & a_1d_1 + a_2d_2 + \dots + a_nd_n \\ b_1c_1 + b_2c_2 + \dots + b_nc_n & b_1d_1 + b_2d_2 + \dots + b_nd_n \end{vmatrix} =$$
$$= \sum_{1 \le i < k \le n} \begin{vmatrix} a_i & a_k \\ b_i & b_k \end{vmatrix} \begin{vmatrix} c_i & d_i \\ c_k & d_k \end{vmatrix}.$$

2. Так как $\begin{pmatrix} a_1c_1 + b_1d_1 & a_1c_2 + b_1d_2 & \dots & a_1c_n + b_1d_n \\ a_2c_1 + b_2d_1 & a_2c_2 + b_2d_2 & \dots & a_2c_n + b_2d_n \\ \dots & \dots & \dots & \dots \\ a_nc_1 + b_nd_1 & a_nc_2 + b_nd_2 & \dots & a_nc_n + b_nd_n \end{pmatrix} =$

$$= \begin{pmatrix} a_1 & b_1 \\ a_2 & b_2 \\ \cdot & \cdot \\ \cdot & \cdot \\ a_n & b_n \end{pmatrix} \cdot \begin{pmatrix} c_1 & c_2 & \dots & c_n \\ d_1 & d_2 & \dots & d_n \end{pmatrix},$$

при $n > 2$ справедливо равенство

$$\begin{vmatrix} a_1c_1 + b_1d_1 & a_1c_2 + b_1d_2 & \dots & a_1c_n + b_1d_n \\ a_2c_1 + b_2d_1 & a_2c_2 + b_2d_2 & \dots & a_2c_n + b_2d_n \\ \dots & \dots & \dots & \dots \\ a_nc_1 + b_nd_1 & a_nc_2 + b_nd_2 & \dots & a_nc_n + b_nd_n \end{vmatrix} = 0.$$

Формула Бине–Коши дает возможность выразить миноры произведения двух матриц через миноры матриц-сомножителей. Пусть

$\boldsymbol{C} = \boldsymbol{A} \cdot \boldsymbol{B}$, где $\boldsymbol{A} = (a_{ij})_{m\times n}$, $\boldsymbol{B} = (b_{jk})_{n\times \ell}$. Рассмотрим произвольный минор матрицы $\boldsymbol{C}$: $\boldsymbol{C}\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix}$, где $1 \le i_1 < i_2 < \ldots < i_p \le m$, $1 \le j_1 < j_2 < \ldots < j_p \le \ell$.

Матрица, составленная из элементов этого минора, является произведением подматрицы матрицы $\boldsymbol{A}$ и подматрицы матрицы $\boldsymbol{B}$:

$$\begin{pmatrix} a_{i_1 1} & a_{i_1 2} & \dots & a_{i_1 n} \\ \cdots & \cdots & \cdots & \cdots \\ a_{i_p 1} & a_{i_p 2} & \dots & a_{i_p n} \end{pmatrix} \cdot \begin{pmatrix} b_{1j_1} & \cdots & b_{1j_p} \\ b_{2j_1} & \cdots & b_{2j_p} \\ \cdots & \cdots & \cdots \\ b_{nj_1} & \cdots & b_{nj_p} \end{pmatrix}.$$

Применяя формулу Бине–Коши, получаем

$$\boldsymbol{C}\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} = \tag{5}$$

$$= \sum_{1 \le k_1 < k_2 < \ldots < k_m \le n} \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ k_1 & k_2 & \dots & k_p \end{pmatrix} \times$$

$$\times \boldsymbol{B}\begin{pmatrix} k_1 & k_2 & \dots & k_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix}.$$

При $p = 1$ получается формула для элемента произведения матриц.

При $p > n$ (в случае, когда $m, \ell > n$) $\boldsymbol{C}\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix} = 0$.

§ 1.2. О решениях матричных уравнений

В этом параграфе рассматриваются матрицы над полем. Применяя определение ранга матрицы по минорам, с помощью формулы (5) получаем новое доказательство известного утверждения.

Теорема 2. *Ранг произведения двух матриц не превосходит ранга каждой из них.*

Для матричных уравнений вида $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{C}$ справедлива следующая

Теорема 3. *Пусть матрицы $\boldsymbol{A}_{m\times n}$ и $\boldsymbol{C}_{m\times \ell}$ таковы, что матричное уравнение $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{C}$ имеет решение. Тогда это уравнение* *имеет решение* $\boldsymbol{X}_0$ *наименьшего возможного ранга* (*равного* $\mathsf{r}(\boldsymbol{C})$), *причем указанное решение можно выбрать в виде* $\boldsymbol{X}_0 = \boldsymbol{L} \cdot \boldsymbol{C}$ *для некоторой матрицы* $\boldsymbol{L}_{m\times n}$.

Доказательство. Ранг любого решения матричного уравнения $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{C}$ не может быть меньше, чем $\mathsf{r}(\boldsymbol{C})$. Пусть $r = \mathsf{r}(\boldsymbol{C})$. Используя определение ранга матрицы по столбцам, предположим, что первые r столбцов матрицы $\boldsymbol{C}$ линейно независимы, а остальные линейно выражаются через них: $\boldsymbol{C}_{\bullet j} = \sum_{k=1}^{r} t_{jk}\boldsymbol{C}_{\bullet k}$ $(j = r+1, \ldots, \ell)$.

Рассмотрим любое решение $\boldsymbol{X}$ матричного уравнения $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{C}$. Тогда для любого столбца $\boldsymbol{X}_{\bullet k}$ этого решения справедливо равенство $\boldsymbol{A} \cdot \boldsymbol{X}_{\bullet k} = \boldsymbol{C}_{\bullet k}$.

Возьмем первые r столбцов матрицы $\boldsymbol{X}$. Они линейно независимы, так как через них линейно выражаются первые r столбцов матрицы $\boldsymbol{C}$. Положим $\tilde{\boldsymbol{X}}_{\bullet j} = \sum_{k=1}^{r} t_{jk}\boldsymbol{X}_{\bullet k}$ при $j = r+1, \ldots, \ell$. Добавив полученные столбцы к первым r столбцам матрицы $\boldsymbol{X}$, получим матрицу $\boldsymbol{X}_0$ ранга r, которая будет решением матричного уравнения $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{C}$.

Покажем, что решение $\boldsymbol{X}_0$ минимального ранга матричного уравнения $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{C}$ всегда можно представить в виде $\boldsymbol{X}_0 = \boldsymbol{L} \cdot \boldsymbol{C}$ для некоторой матрицы $\boldsymbol{L}_{m\times n}$.

Из равенства $\boldsymbol{A} \cdot \boldsymbol{X}_0 = \boldsymbol{C}$ следует, что строки матрицы $\boldsymbol{C}$ являются линейными комбинациями строк матрицы $\boldsymbol{X}_0$. Так как ранги (по строкам) матриц $\boldsymbol{C}$ и $\boldsymbol{X}_0$ совпадают, линейные оболочки систем строк этих матриц также равны.

Отсюда следует, что строки матрицы $\boldsymbol{X}_0$ являются линейными комбинациями строк матрицы $\boldsymbol{C}$ и поэтому $\boldsymbol{X}_0 = \boldsymbol{L} \cdot \boldsymbol{C}$ для некоторой матрицы $\boldsymbol{L}_{m\times n}$. Теорема доказана.

Для матричных уравнений вида $\boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{B} = \boldsymbol{C}$ имеет место

Теорема 4. *Матричное уравнение вида* $\boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{B} = \boldsymbol{C}$, *где* $\boldsymbol{A}$,$\boldsymbol{B}$,$\boldsymbol{C}$ – *известные матрицы, имеет решение тогда и только тогда, когда одновременно имеют решения матричные уравнения* $\boldsymbol{A} \cdot \boldsymbol{Y} = \boldsymbol{C}$ *и* $\boldsymbol{Z} \cdot \boldsymbol{B} = \boldsymbol{C}$.

Доказательство. Очевидно, что если матричное уравнение $\boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{B} = \boldsymbol{C}$ имеет решение $\boldsymbol{X}$, то и уравнения $\boldsymbol{A} \cdot \boldsymbol{Y} = \boldsymbol{C}$ и $\boldsymbol{Z} \cdot \boldsymbol{B} = \boldsymbol{C}$ имеют решения $\boldsymbol{Y} = \boldsymbol{X} \cdot \boldsymbol{B}$ и $\boldsymbol{Z} = \boldsymbol{A} \cdot \boldsymbol{X}$.

Обратно, пусть уравнения $\boldsymbol{A}\cdot\boldsymbol{Y} = \boldsymbol{C}$ и $\boldsymbol{Z}\cdot\boldsymbol{B} = \boldsymbol{C}$ имеют решения $\boldsymbol{Y}$ и $\boldsymbol{Z}$. Тогда первое из указанных уравнений имеет решение $\boldsymbol{Y}_0$ минимального ранга $\mathsf{r}(\boldsymbol{C})$, которое в силу теоремы 1 может быть представлено в виде $\boldsymbol{Y}_0 = \boldsymbol{L}\cdot\boldsymbol{C}$. Следовательно, $\boldsymbol{C} = \boldsymbol{A}\cdot\boldsymbol{Y}_0 = \boldsymbol{A}\cdot\boldsymbol{L}\cdot\boldsymbol{C} = \boldsymbol{A}\cdot\boldsymbol{L}\cdot\boldsymbol{Z}\cdot\boldsymbol{B}$. Теперь матрица $\boldsymbol{L}\cdot\boldsymbol{Z}$ является решением матричного уравнения $\boldsymbol{A}\cdot\boldsymbol{X}\cdot\boldsymbol{B} = \boldsymbol{C}$.

§ 1.3. Ассоциированная матрица

В этой теме рассматриваются матрицы над полем.

Пусть $\boldsymbol{A} = (a_{ij})_{n\times n}$ – квадратная матрица порядка n. Рассмотрим всевозможные миноры порядка p ($1 \le p \le n$) матрицы $\boldsymbol{A}$. Их количество равно N_p^2, где $N_p = C_n^p$. Составим из миноров $\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix}$ ($1 \le i_1 < i_2 < \ldots < i_p \le n$ и $1 \le j_1 < j_2 < \ldots < j_p \le n$) квадратную матрицу порядка N_p.

Для этого линейно упорядочим определенным способом (например, лексикографически) все сочетания из n индексов $1, 2, \ldots, n$ по p. Этот порядок фиксируется для всех квадратных матриц порядка n. Если сочетания $1 \le i_1 < i_2 < \ldots < i_p \le n$ и $1 \le j_1 < j_2 < \ldots < j_p \le n$ имеют в указанном порядке номера α и β, то положим $\breve{a}_{\alpha\beta} = \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \dots & i_p \\ j_1 & j_2 & \dots & j_p \end{pmatrix}$.

Квадратная матрица $\breve{\boldsymbol{A}}_p = (\breve{a}_{\alpha\beta})_{N_p\times N_p}$ порядка N_p называется p-й *ассоциированной матрицей* для матрицы $\boldsymbol{A} = (a_{ij})_{n\times n}$.

Здесь p может принимать значения $1, 2, \ldots, n$. При этом $\breve{\boldsymbol{A}}_1 = \boldsymbol{A}$, а матрица $\breve{\boldsymbol{A}}_n$ состоит из одного элемента, равного $|\boldsymbol{A}|$.

Рассмотрим пример.

Пусть $\boldsymbol{A} = \begin{pmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{pmatrix}$. Построим матрицу $\breve{\boldsymbol{A}}_2$.

Для этого упорядочим сочетания из 4 по 2 следующим образом: (1 2), (1 3), (1 4), (2 3), (2 4), (3 4). Тогда $\breve{\boldsymbol{A}}_2 =$

$$
= \begin{pmatrix}
A\begin{pmatrix} 1 & 2 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 2 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 2 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 2 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix} \\
A\begin{pmatrix} 1 & 3 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 1 & 3 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 3 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 3 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 3 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 3 \\ 3 & 4 \end{pmatrix} \\
A\begin{pmatrix} 1 & 4 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 1 & 4 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 4 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 1 & 4 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 1 & 4 \\ 3 & 4 \end{pmatrix} \\
A\begin{pmatrix} 2 & 3 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 2 & 3 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 2 & 3 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 2 & 3 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 2 & 3 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 2 & 3 \\ 3 & 4 \end{pmatrix} \\
A\begin{pmatrix} 2 & 4 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 2 & 4 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 2 & 4 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 2 & 4 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 2 & 4 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 2 & 4 \\ 3 & 4 \end{pmatrix} \\
A\begin{pmatrix} 3 & 4 \\ 1 & 2 \end{pmatrix} & A\begin{pmatrix} 3 & 4 \\ 1 & 3 \end{pmatrix} & A\begin{pmatrix} 3 & 4 \\ 1 & 4 \end{pmatrix} & A\begin{pmatrix} 3 & 4 \\ 2 & 3 \end{pmatrix} & A\begin{pmatrix} 3 & 4 \\ 2 & 4 \end{pmatrix} & A\begin{pmatrix} 3 & 4 \\ 3 & 4 \end{pmatrix}
\end{pmatrix}.
$$

Предложение 1. *Матрица p-я ассоциированная с единичной матрицей $\boldsymbol{E}_n$ порядка n есть единичная матрица $\boldsymbol{E}_{N_p}$ порядка $N_p = C_n^p$.*

Доказательство. Минор называется *главным*, если в нем номера строк и столбцов совпадают. В любой ассоциированной матрице $\breve{\boldsymbol{A}}_p$ все главные миноры порядка p матрицы $\boldsymbol{A}$ образуют главную диагональ. Легко видеть, что в единичной матрице любой главный минор равен 1, а любой минор, не являющийся главным, равен 0. Отсюда следует требуемое утверждение.

Теорема 5. *Пусть $\boldsymbol{C} = \boldsymbol{A} \cdot \boldsymbol{B}$ (все матрицы квадратные порядка n). Тогда для любого $p \in \{1, 2, \ldots, n\}$ справедливо $\breve{\boldsymbol{C}}_p = \breve{\boldsymbol{A}}_p \cdot \breve{\boldsymbol{B}}_p$.*

Доказательство. По формуле (5) из § 1.1 имеем для любых $1 \le i_1 < i_2 < \ldots < i_p \le n$, $1 \le j_1 < j_2 < \ldots < j_p \le n$

$$
\boldsymbol{C}\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ j_1 & j_2 & \ldots & j_p \end{pmatrix} =
$$
$$
= \sum\nolimits_{1 \le k_1 < k_2 < \ldots < k_p \le n} \boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix} \boldsymbol{B}\begin{pmatrix} k_1 & k_2 & \ldots & k_p \\ j_1 & j_2 & \ldots & j_p \end{pmatrix}.
$$

В обозначениях этого параграфа последнее равенство можно записать так:

$$
\breve{c}_{\alpha\beta} = \sum_{\lambda=1}^{N_p} \breve{a}_{\alpha\lambda} \breve{b}_{\lambda\beta},
$$

где α, β, λ — номера сочетаний $1 \le i_1 < i_2 < \ldots < i_p \le n$, $1 \le$ $\le j_1 < j_2 < \ldots < j_p \le n$, $1 \le k_1 < k_2 < \ldots < k_p \le n$ соответственно. Отсюда вытекает требуемое равенство.

Из предложения и теоремы получается

Следствие 1. *Если квадратная матрица $\boldsymbol{A}$ порядка n обратима и матрица $\boldsymbol{B} = \boldsymbol{A}^{-1}$, то для любого $p \in \{1, 2, \ldots, n\}$ справедливо $\breve{\boldsymbol{B}}_p = (\breve{\boldsymbol{A}}_p)^{-1}$.*

Для доказательства заметим, что $\boldsymbol{A}\cdot\boldsymbol{B}=\boldsymbol{E}_n$, откуда в силу теоремы и предложения выводим $\breve{\boldsymbol{A}}_p\cdot\breve{\boldsymbol{B}}_p=\boldsymbol{E}_{N_p}$.

§ 1.4. Выражение миноров обратной матрицы через миноры исходной матрицы

Пусть квадратная матрица $\boldsymbol{A}$ порядка n обратима и матрица $\boldsymbol{B}=\boldsymbol{A}^{-1}$. Тогда для любого $p\in\{1,2,\ldots,n\}$ имеет место формула для минора обратной матрицы (используем формулу (2) из § 1.1 для алгебраического дополнения к минору):

$$\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}=\frac{\mathsf{ad}\boldsymbol{A}\begin{pmatrix} k_1 & k_2 & \ldots & k_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix}}{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & n \\ 1 & 2 & \ldots & n \end{pmatrix}}. \tag{1}$$

Доказательство. Из $\boldsymbol{A}\cdot\boldsymbol{B}=\boldsymbol{E}_n$ следует $\breve{\boldsymbol{A}}_p\cdot\breve{\boldsymbol{B}}_p=\boldsymbol{E}_{N_p}$, или

$$\sum_{\lambda=1}^{N_p}\breve{a}_{\alpha\lambda}\breve{b}_{\lambda\beta}=\begin{cases} 1, & \alpha=\beta; \\ 0, & \alpha\neq\beta. \end{cases}$$

Эти равенства могут быть записаны следующим образом:

$$\sum_{1\le i_1<i_2<\ldots<i_p\le n}\boldsymbol{A}\begin{pmatrix} j_1 & j_2 & \ldots & j_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix}\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \ldots & i_p \\ k_1 & k_2 & \ldots & k_p \end{pmatrix}= \tag{2}$$

$$=\begin{cases} 1, & \sum_{t=1}^{p}(j_t-k_t)^2=0; \\ 0, & \sum_{t=1}^{p}(j_t-k_t)^2>0 \end{cases}$$

($1\le j_1<j_2<\ldots<j_p\le n$ и $1\le k_1<k_2<\ldots<k_p\le n$.)

С другой стороны, по теореме Лапласа, примененной к определителю матрицы $\boldsymbol{A}$, для любых $1\le j_1<j_2<\ldots<j_p\le n$ имеем

$$\sum_{1\le i_1<i_2<\ldots<i_p\le n}\boldsymbol{A}\begin{pmatrix} j_1 & j_2 & \ldots & j_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix}\mathsf{ad}\boldsymbol{A}\begin{pmatrix} j_1 & j_2 & \ldots & j_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix}=|\boldsymbol{A}|.$$

Если взять сочетание $1\le k_1<k_2<\ldots<k_p\le n$, отличное от $1\le j_1<j_2<\ldots<j_p\le n$, то

$$\sum_{1 \le i_1 < i_2 < \ldots < i_p \le n} \boldsymbol{A}\begin{pmatrix} j_1 & j_2 & \ldots & j_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} \mathsf{ad}\boldsymbol{A}\begin{pmatrix} k_1 & k_2 & \ldots & k_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} = 0.$$

Это утверждение следует из разложения Лапласа, примененного к матрице, полученной из $\boldsymbol{A}$ заменой строк с номерами $1 \le j_1 < < j_2 < \ldots < j_p \le n$ на строки с номерами $1 \le k_1 < k_2 < \ldots < k_p \le \le n$: в ней появятся одинаковые строки и определитель ее будет равен нулю.

Следовательно,

$$\sum_{1 \le i_1 < i_2 < \ldots < i_p \le n} \boldsymbol{A}\begin{pmatrix} j_1 & j_2 & \ldots & j_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} = \mathsf{ad}\boldsymbol{A}\begin{pmatrix} j_1 & j_2 & \ldots & j_p \\ i_1 & i_2 & \ldots & i_p \end{pmatrix} = \\ = \begin{cases} |\boldsymbol{A}|, \ \sum_{t=1}^{p}(j_t - k_t)^2 = 0; \\ 0, \ \sum_{t=1}^{p}(j_t - k_t)^2 > 0. \end{cases}$$

Сопоставляя полученные формулы с формулами (2), получаем требуемое, так как из равенства $\breve{\boldsymbol{A}} \cdot \boldsymbol{X} = \boldsymbol{E}_{N_p}$ следует $\boldsymbol{X} = \breve{\boldsymbol{A}}^{-1}$.

§ 1.5. Псевдообратная матрица

В этом параграфе рассматриваются матрицы над полем комплексных чисел.

Хорошо известно, что обратная матрица существует только у квадратной невырожденной матрицы. Для прямоугольной или квадратной вырожденной матрицы обратная матрица не существует.

Оказывается возможным определить для любой матрицы $\boldsymbol{A}$ псевдообратную матрицу $\boldsymbol{A}^+$, которая обладает некоторыми свойствами обратной матрицы, имеет важные применения к системам линейных уравнений, а в случае обратимой матрицы $\boldsymbol{A}$ совпадает с $\boldsymbol{A}^{-1}$.

Для матрицы $\boldsymbol{A}$ через $\boldsymbol{A}^*$ обозначим сопряженную матрицу (транспонированную к $\boldsymbol{A}$, в которой каждый элемент заменен комплексно сопряженным числом). Если произведение $\boldsymbol{A} \cdot \boldsymbol{B}$ матриц определено, то

$$(\boldsymbol{A} \cdot \boldsymbol{B})^* = \boldsymbol{B}^* \cdot \boldsymbol{A}^*.$$

Рассмотрим матричное уравнение $\boldsymbol{A} \cdot \boldsymbol{X} \cdot \boldsymbol{A} = \boldsymbol{A}$, где $\boldsymbol{A}$ – известная матрица размеров $m \times n$. Если $\boldsymbol{A}$ – квадратная невырожденная матрица, то уравнение имеет единственное решение $\boldsymbol{X} = \boldsymbol{A}^{-1}$.

Если $\boldsymbol{A}$ не является обратимой матрицей, то уравнение имеет решение размеров $n \times m$. Будет показано, что такое решение существует и единственно при условии, что его строки и столбцы линейно выражаются через строки и столбцы сопряженной матрицы $\boldsymbol{A}^* = \overline{\boldsymbol{A}}^\top$. Именно это решение называется псевдообратной матрицей к матрице $\boldsymbol{A}$ и обозначается через $\boldsymbol{A}^+$.

Матрица $\boldsymbol{A}^+$ размеров $n \times m$ называется *псевдообратной* матрицей к матрице $\boldsymbol{A}$ размеров $m \times n$, если выполняются следующие условия:

$$\boldsymbol{A} \cdot \boldsymbol{A}^+ \cdot \boldsymbol{A} = \boldsymbol{A}; \tag{1}$$

$$\boldsymbol{A}^+ = \boldsymbol{U} \cdot \boldsymbol{A}^* = \boldsymbol{A}^* \cdot \boldsymbol{V}, \text{ где } \boldsymbol{U}, \boldsymbol{V} \text{ – некоторые матрицы.} \tag{2}$$

Первое из равенств (2) показывает, что строки матрицы $\boldsymbol{A}^+$ линейно выражаются через строки матрицы $\boldsymbol{A}^*$, а второе обеспечивает аналогичное утверждение для столбцов.

В силу теоремы 4 (см. § 1.2) условие (2) можно заменить условием $\boldsymbol{A}^+ = \boldsymbol{A}^* \cdot \boldsymbol{W} \cdot \boldsymbol{A}^*$ для некоторой матрицы $\boldsymbol{W}$. Для построения псевдообратной матрицы важную роль играет следующее понятие.

Пусть $\boldsymbol{A}$ – матрица размеров $m \times n$ ранга $r > 0$. *Скелетным разложением* матрицы $\boldsymbol{A}$ называется ее представление в виде произведения $\boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{C}$ матрицы $\boldsymbol{B}$ размеров $m \times r$ на матрицу $\boldsymbol{C}$ размеров $r \times n$.

Скелетное разложение имеется у любой ненулевой матрицы. Для его нахождения достаточно взять r линейно независимых столбцов матрицы $\boldsymbol{A}$ (это будет матрица $\boldsymbol{B}$) и линейно выразить через них все столбцы матрицы $\boldsymbol{A}$. Матрицу, составленную из столбцов коэффициентов полученных линейных комбинаций, можно взять в качестве матрицы $\boldsymbol{C}$. Скелетное разложение определено неоднозначно.

Так как ранг произведения матриц не превосходит ранга каждого из сомножителей, в любом скелетном разложении $\boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{C}$ матрицы ранга $r > 0$ имеет место $\mathsf{r}(\boldsymbol{B}) = r$ и $\mathsf{r}(\boldsymbol{C}) = r$.

Рассмотрим пример. Найти скелетное разложение матрицы

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix}.$$

Вычислим ранг матрицы $\boldsymbol{A}$ с помощью элементарных преобразований строк: $\boldsymbol{A} \longrightarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -3 \\ 0 & -2 & -4 & -6 \end{pmatrix} \longrightarrow$

$$\longrightarrow \begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & -1 & -2 & -3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \longrightarrow \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{pmatrix}.$$

Мы видим, что $\mathsf{r}(\boldsymbol{A}) = 2$ и первые два столбца матрицы $\boldsymbol{A}$ линейно независимы. Поэтому в качестве скелетного разложения матрицы $\boldsymbol{A}$ можно взять представление

$$\boldsymbol{A} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{pmatrix}.$$

Предложение 2. *Пусть матрицы $\boldsymbol{B}$ размеров $m \times r$ и $\boldsymbol{C}$ размеров $r \times n$ имеют ранг r. Тогда матрицы $\boldsymbol{B}^* \cdot \boldsymbol{B}$ и $\boldsymbol{C} \cdot \boldsymbol{C}^*$ являются невырожденными, т. е. $|\boldsymbol{B}^* \cdot \boldsymbol{B}| \neq 0$ и $|\boldsymbol{C} \cdot \boldsymbol{C}^*| \neq 0$.*

Доказательство. Матрицу $\boldsymbol{B}^* \cdot \boldsymbol{B}$ можно рассматривать как матрицу Грама системы столбцов матрицы $\boldsymbol{B}$ в арифметическом унитарном пространстве столбцов $\mathbb{C}^{m\times 1}$. Так как столбцы матрицы $\boldsymbol{B}$ линейно независимы в силу условия $\mathsf{r}(\boldsymbol{B}) = r$, определитель Грама системы этих столбцов не равен нулю. Следовательно, $|\boldsymbol{B}^* \cdot \boldsymbol{B}| \neq 0$. Для матрицы $\boldsymbol{C} \cdot \boldsymbol{C}^*$ доказательство проводится аналогично, только вместо столбцов следует рассматривать строки.

Докажем существование матрицы $\boldsymbol{A}^+$. Пусть $\boldsymbol{A}$ – произвольная матрица размеров $m \times n$. Если $\boldsymbol{A}$ – нулевая матрица, то $\boldsymbol{A}^+$ – тоже нулевая. Предположим, что $\mathsf{r}(\boldsymbol{A}) = r > 0$. Рассмотрим скелетное разложение $\boldsymbol{B} \cdot \boldsymbol{C}$ матрицы $\boldsymbol{A}$. Найдем матрицу $\boldsymbol{B}^+$.

По определению $\boldsymbol{B} \cdot \boldsymbol{B}^+ \cdot \boldsymbol{B} = \boldsymbol{B}$ и $\boldsymbol{B}^+ = \widehat{\boldsymbol{U}} \cdot \boldsymbol{B}^*$. Подставим в первое равенство $\boldsymbol{B}^+$ из второго: $\boldsymbol{B} \cdot \widehat{\boldsymbol{U}} \cdot \boldsymbol{B}^* \cdot \boldsymbol{B} = \boldsymbol{B}$, и умножим обе части последнего равенства на $\boldsymbol{B}^*$ слева: $\boldsymbol{B}^* \cdot \boldsymbol{B} \cdot \widehat{\boldsymbol{U}} \cdot \boldsymbol{B}^* \cdot \boldsymbol{B} = \boldsymbol{B}^* \cdot \boldsymbol{B}$.

Так как матрица $\boldsymbol{B}^* \cdot \boldsymbol{B}$ невырожденная и потому обратимая, находим матрицу $\widehat{\boldsymbol{U}} = (\boldsymbol{B}^* \cdot \boldsymbol{B})^{-1}$.

Следовательно, получаем формулу для псевдообратной матрицы к матрице размеров $m \times r$ ранга r:

$$\boldsymbol{B}^{+} = (\boldsymbol{B}^{*} \cdot \boldsymbol{B})^{-1} \cdot \boldsymbol{B}^{*}. \tag{3}$$

Совершенно аналогично найдем формулу для псевдообратной матрицы к матрице размеров $r \times n$ ранга r

$$\boldsymbol{C}^{+} = \boldsymbol{C}^{*} \cdot (\boldsymbol{C} \cdot \boldsymbol{C}^{*})^{-1}. \tag{4}$$

Покажем, что матрица

$$\boldsymbol{A}^{+} = \boldsymbol{C}^{+} \cdot \boldsymbol{B}^{+} = \boldsymbol{C}^{*} \cdot (\boldsymbol{C} \cdot \boldsymbol{C}^{*})^{-1} \cdot (\boldsymbol{B}^{*} \cdot \boldsymbol{B})^{-1} \cdot \boldsymbol{B}^{*} \tag{5}$$

удовлетворяет условиям (1) и (2), т. е. является псевдообратной для матрицы $\boldsymbol{A}$.

Имеем $\boldsymbol{A} \cdot \boldsymbol{A}^{+} \cdot \boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{C} \cdot \boldsymbol{C}^{*} \cdot (\boldsymbol{C} \cdot \boldsymbol{C}^{*})^{-1} \cdot (\boldsymbol{B}^{*} \cdot \boldsymbol{B})^{-1} \cdot \boldsymbol{B}^{*} \cdot \boldsymbol{B} \cdot \boldsymbol{C} = \boldsymbol{B} \cdot \boldsymbol{C} = \boldsymbol{A}$.

Далее, $\boldsymbol{A}^{*} = \boldsymbol{C}^{*} \cdot \boldsymbol{B}^{*}$. Обозначим для краткости $\boldsymbol{K} = (\boldsymbol{C} \cdot \boldsymbol{C}^{*})^{-1} \cdot (\boldsymbol{B}^{*} \cdot \boldsymbol{B})^{-1}$. Тогда $\boldsymbol{A}^{+} = \boldsymbol{C}^{*} \cdot \boldsymbol{K} \cdot \boldsymbol{B}^{*} = \boldsymbol{C}^{*} \cdot \boldsymbol{K} \cdot (\boldsymbol{C} \cdot \boldsymbol{C}^{*})^{-1} \cdot \boldsymbol{C} \cdot \boldsymbol{C}^{*} \cdot \boldsymbol{B}^{*} = \boldsymbol{U} \cdot \boldsymbol{C}^{*} \cdot \boldsymbol{B}^{*} = \boldsymbol{U} \cdot \boldsymbol{A}^{*}$, где $\boldsymbol{U} = \boldsymbol{C}^{*} \cdot \boldsymbol{K} \cdot (\boldsymbol{C} \cdot \boldsymbol{C}^{*})^{-1} \cdot \boldsymbol{C}$.

Аналогично $\boldsymbol{A}^{+} = \boldsymbol{C}^{*} \cdot \boldsymbol{K} \cdot \boldsymbol{B}^{*} = \boldsymbol{C}^{*} \cdot \boldsymbol{B}^{*} \cdot \boldsymbol{B} \cdot (\boldsymbol{B}^{*} \cdot \boldsymbol{B})^{-1} \cdot \boldsymbol{K} \cdot \boldsymbol{B}^{*} = \boldsymbol{C}^{*} \cdot \boldsymbol{B}^{*} \cdot \boldsymbol{V} = \boldsymbol{A}^{*} \cdot \boldsymbol{V}$, где $\boldsymbol{V} = \boldsymbol{B} \cdot (\boldsymbol{B}^{*} \cdot \boldsymbol{B})^{-1} \cdot \boldsymbol{K} \cdot \boldsymbol{B}^{*}$.

Доказательство единственности псевдообратной матрицы опирается на следующую лемму.

Лемма 1. *Для любой матрицы $\boldsymbol{D}$ размеров $m \times n$ из равенства $\boldsymbol{D}^{*} \cdot \boldsymbol{D} = \boldsymbol{O}$ следует $\boldsymbol{D} = \boldsymbol{O}$.*

Доказательство. Матрицу $\boldsymbol{D}^{*} \cdot \boldsymbol{D}$ можно рассматривать как матрицу Грама системы столбцов матрицы $\boldsymbol{D}$ в арифметическом унитарном пространстве столбцов $\mathbb{C}^{m \times 1}$. Матрица Грама ненулевой системы векторов имеет ненулевые элементы на главной диагонали.

Предположим, что для некоторой матрицы $\boldsymbol{A}$ существуют две псевдообратные матрицы $\boldsymbol{A}_1^{+}$ и $\boldsymbol{A}_2^{+}$. Тогда $\boldsymbol{A}\boldsymbol{A}_1^{+}\boldsymbol{A} = \boldsymbol{A}$, $\boldsymbol{A}\boldsymbol{A}_2^{+}\boldsymbol{A} = \boldsymbol{A}$, $\boldsymbol{A}_1^{+} = \boldsymbol{U}_1 \cdot \boldsymbol{A}^{*} = \boldsymbol{A}^{*} \cdot \boldsymbol{V}_1$, $\boldsymbol{A}_2^{+} = \boldsymbol{U}_2 \cdot \boldsymbol{A}^{*} = \boldsymbol{A}^{*} \cdot \boldsymbol{V}_2$ для некоторых матриц $\boldsymbol{U}_1$, $\boldsymbol{U}_2$, $\boldsymbol{V}_1$, $\boldsymbol{V}_2$. Положим $\boldsymbol{D} = \boldsymbol{A}_2^{+} - \boldsymbol{A}_1^{+}$, $\boldsymbol{U} = \boldsymbol{U}_2 - \boldsymbol{U}_1$, $\boldsymbol{V} = \boldsymbol{V}_2 - \boldsymbol{V}_1$. Тогда $\boldsymbol{A} \cdot \boldsymbol{D} \cdot \boldsymbol{A} = \boldsymbol{O}$ и $\boldsymbol{D} = \boldsymbol{U} \cdot \boldsymbol{A}^{*} = \boldsymbol{A}^{*} \cdot \boldsymbol{V}$. Отсюда $(\boldsymbol{D} \cdot \boldsymbol{A})^{*} \cdot \boldsymbol{D} \cdot \boldsymbol{A} = \boldsymbol{A}^{*} \cdot \boldsymbol{D}^{*} \cdot \boldsymbol{D} \cdot \boldsymbol{A} = \boldsymbol{A}^{*} \cdot (\boldsymbol{A}^{*} \cdot \boldsymbol{V})^{*} \cdot \boldsymbol{D} \cdot \boldsymbol{A} = \boldsymbol{A}^{*} \cdot \boldsymbol{V}^{*} \cdot \boldsymbol{A} \cdot \boldsymbol{D} \cdot \boldsymbol{A} = \boldsymbol{O}$. По лемме $\boldsymbol{D} \cdot \boldsymbol{A} = \boldsymbol{O}$. Далее, $\boldsymbol{D} \cdot \boldsymbol{D}^{*} = \boldsymbol{D} \cdot (\boldsymbol{U} \cdot \boldsymbol{A}^{*})^{*} = \boldsymbol{D} \cdot \boldsymbol{A} \cdot \boldsymbol{U}^{*} =$

$= \boldsymbol{O}$. Снова по лемме $\boldsymbol{D} = \boldsymbol{O}$ и $\boldsymbol{A}_2^+ - \boldsymbol{A}_1^+ = \boldsymbol{O}$, т. е. $\boldsymbol{A}_2^+ = \boldsymbol{A}_1^+$.

Найдем псевдообратную для матрицы $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \end{pmatrix}$.

Воспользуемся ее скелетным разложением $\boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{C}$, где $\boldsymbol{B} = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{pmatrix}$. Имеем $\boldsymbol{B}^* \cdot \boldsymbol{B} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 4 \end{pmatrix} = \begin{pmatrix} 14 & 20 \\ 20 & 29 \end{pmatrix}$ и $(\boldsymbol{B}^* \cdot \boldsymbol{B})^{-1} \cdot \boldsymbol{B}^* = \frac{1}{6} \begin{pmatrix} 29 & -20 \\ -20 & 14 \end{pmatrix} \cdot \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} = \frac{1}{6} \begin{pmatrix} -11 & -2 & 7 \\ 8 & 2 & -4 \end{pmatrix}$.

Аналогично $\boldsymbol{C} \cdot \boldsymbol{C}^* = \begin{pmatrix} 1 & 0 & -1 & -2 \\ 0 & 1 & 2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 2 \\ -2 & 3 \end{pmatrix} = \begin{pmatrix} 6 & -8 \\ -8 & 14 \end{pmatrix}$ и $(\boldsymbol{C} \cdot \boldsymbol{C}^*)^{-1} = \frac{1}{20} \begin{pmatrix} 14 & 8 \\ 8 & 6 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 7 & 4 \\ 4 & 3 \end{pmatrix}$,

$$\boldsymbol{C}^* \cdot (\boldsymbol{C} \cdot \boldsymbol{C}^*)^{-1} = \frac{1}{10} \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ -1 & 2 \\ -2 & 3 \end{pmatrix} \cdot \begin{pmatrix} 7 & 4 \\ 4 & 3 \end{pmatrix} = \frac{1}{10} \begin{pmatrix} 7 & 4 \\ 4 & 3 \\ 1 & 2 \\ -2 & 1 \end{pmatrix}.$$

По формуле (5) получаем $\boldsymbol{A}^+ = \boldsymbol{C}^* \cdot (\boldsymbol{C} \cdot \boldsymbol{C}^*)^{-1} \cdot (\boldsymbol{B}^* \cdot \boldsymbol{B})^{-1} \cdot \boldsymbol{B}^* = \frac{1}{10} \begin{pmatrix} 7 & 4 \\ 4 & 3 \\ 1 & 2 \\ -2 & 1 \end{pmatrix} \cdot \frac{1}{6} \begin{pmatrix} -11 & -2 & 7 \\ 8 & 2 & -4 \end{pmatrix} = \frac{1}{60} \begin{pmatrix} -45 & -6 & 33 \\ -20 & -2 & 16 \\ 5 & 2 & -1 \\ 30 & 6 & -18 \end{pmatrix}$.

Из утверждения о существовании и единственности псевдообратной матрицы и ее построения с помощью скелетного разложения сразу получается следующее утверждение.

Предложение 3. *Для любого скелетного разложения* $\boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{C}$ *матрицы* $\boldsymbol{A}$ *справедливы формулы (5), в частности* $\boldsymbol{A}^+ = \boldsymbol{C}^+ \cdot \boldsymbol{B}^+$.

Если разложение $\boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{C}$ не является скелетным для матрицы $\boldsymbol{A}$, то равенство $\boldsymbol{A}^+ = \boldsymbol{C}^+ \cdot \boldsymbol{B}^+$ может не выполняться. Читателю предлагается проверить это для разложения $\boldsymbol{A} = (1) = (1\ 0) \cdot \begin{pmatrix} 1 \\ 1 \end{pmatrix}$.

Теорема 6. *Для любой матрицы* $\boldsymbol{A}$ *справедливы следующие утверждения:*

1) $(\boldsymbol{A}^*)^+ = (\boldsymbol{A}^+)^*$*;*

2) $(\boldsymbol{A}^+)^+ = \boldsymbol{A}$*;*

3) $(\boldsymbol{A}\cdot\boldsymbol{A}^+)^* = \boldsymbol{A}\cdot\boldsymbol{A}^+$, $(\boldsymbol{A}\cdot\boldsymbol{A}^+)^2 = \boldsymbol{A}\cdot\boldsymbol{A}^+$*;*

4) $(\boldsymbol{A}^+\cdot\boldsymbol{A})^* = \boldsymbol{A}^+\cdot\boldsymbol{A}$, $(\boldsymbol{A}^+\cdot\boldsymbol{A})^2 = \boldsymbol{A}^+\cdot\boldsymbol{A}$*.*

Доказательство. 1. Возьмем скелетное разложение $\boldsymbol{A} = \boldsymbol{B}\cdot\boldsymbol{C}$. Тогда $\boldsymbol{A}^* = \boldsymbol{C}^*\cdot\boldsymbol{B}^*$ — скелетное разложение матрицы $\boldsymbol{A}^*$. Применяя предложение 3 и формулы (5) и учитывая, что $(\boldsymbol{X}\cdot\boldsymbol{Y})^* = \boldsymbol{Y}^*\cdot\boldsymbol{X}^*$, $(\boldsymbol{X}^*)^{-1} = (\boldsymbol{X}^{-1})^*$, получаем $(\boldsymbol{A}^*)^+ = \boldsymbol{B}^{**}\cdot(\boldsymbol{B}^*\cdot\boldsymbol{B}^{**})^{-1}\cdot(\boldsymbol{C}^{**}\cdot\boldsymbol{C}^*)^{-1}\cdot$ $\cdot\boldsymbol{C}^{**} = (\boldsymbol{C}^*\cdot(\boldsymbol{C}\cdot\boldsymbol{C}^*)^{-1}\cdot(\boldsymbol{B}^*\cdot\boldsymbol{B})^{-1}\cdot\boldsymbol{B}^*)^* = (\boldsymbol{A}^+)^*$.

2. Равенства $\boldsymbol{A}^+ = \boldsymbol{C}^+\cdot\boldsymbol{B}^+$, $\boldsymbol{B}^+ = (\boldsymbol{B}^*\cdot\boldsymbol{B})^{-1}\cdot\boldsymbol{B}^*$, $\boldsymbol{C}^+ =$ $= \boldsymbol{C}^*\cdot(\boldsymbol{C}\cdot\boldsymbol{C}^*)^{-1}$ являются скелетными разложениями. Следовательно, по предложению 3 $(\boldsymbol{A}^+)^+ = (\boldsymbol{B}^+)^+\cdot(\boldsymbol{C}^+)^+$. Далее, с учетом утверждения 1 имеем $(\boldsymbol{B}^+)^+ = (\boldsymbol{B}^*)^+\cdot((\boldsymbol{B}^*\cdot\boldsymbol{B})^{-1})^+ =$
$= (\boldsymbol{B}^+)^*\cdot\boldsymbol{B}^*\cdot\boldsymbol{B} = ((\boldsymbol{B}^*\cdot\boldsymbol{B})^{-1}\cdot\boldsymbol{B}^*)^*\cdot\boldsymbol{B}^*\cdot\boldsymbol{B} = \boldsymbol{B}\cdot(\boldsymbol{B}^*\cdot\boldsymbol{B})^{-1}\boldsymbol{B}^*\cdot\boldsymbol{B} = \boldsymbol{B}$, т. е. $(\boldsymbol{B}^+)^+ = \boldsymbol{B}$. Аналогично получаем $(\boldsymbol{C}^+)^+ = \boldsymbol{C}$. Таким образом, $(\boldsymbol{A}^+)^+ = \boldsymbol{B}\cdot\boldsymbol{C} = \boldsymbol{A}$.

3. Вычислим $\boldsymbol{A}\cdot\boldsymbol{A}^+ = \boldsymbol{B}\cdot\boldsymbol{C}\cdot\boldsymbol{C}^*\cdot(\boldsymbol{C}\cdot\boldsymbol{C}^*)^{-1}\cdot(\boldsymbol{B}^*\cdot\boldsymbol{B})^{-1}\cdot\boldsymbol{B}^* =$ $= \boldsymbol{B}\cdot(\boldsymbol{B}^*\cdot\boldsymbol{B})^{-1}\cdot\boldsymbol{B}^*$.

Теперь ясно, что $(\boldsymbol{A}\cdot\boldsymbol{A}^+)^* = \boldsymbol{A}\cdot\boldsymbol{A}^+$, $(\boldsymbol{A}\cdot\boldsymbol{A}^+)^2 = \boldsymbol{A}\cdot\boldsymbol{A}^+$.

Утверждение 4 проверяется точно так же.

Рассмотрим произвольную, не обязательно совместную систему линейных алгебраических уравнений (СЛАУ) над полем комплексных чисел

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2; \\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots\ \ldots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n = b_m, \end{cases} \tag{6}$$

или в матричной записи

$$\boldsymbol{A}\cdot\boldsymbol{x} = \boldsymbol{b},$$

где $\boldsymbol{A} = (a_{tk})_{m\times n}$, $\boldsymbol{x} = (x_1, x_2, \ldots, x_n)^\top$, $\boldsymbol{b} = (b_1, b_2, \ldots, b_m)^\top$.

Столбец $\boldsymbol{x}^0 = (x_1^0, x_2^0, \ldots, x_n^0)^\top$ называется *наилучшим приближенным решением* СЛАУ (6), если при значениях $x_1 = x_1^0$,

$x_2 = x_2^0, \ldots, x_n = x_n^0$ квадратичное отклонение

$$|\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}|^2 = \sum_{t=1}^{m} \left| b_t - \sum_{k=1}^{n} a_{tk} x_k \right|^2$$

достигает своего наименьшего значения и среди всех столбцов $\boldsymbol{x}$, для которых это отклонение имеет минимальное значение, столбец $\boldsymbol{x}^0$ имеет наименьшую длину как элемент унитарного пространства столбцов $\mathbb{C}_n$, т. е. для этого столбца величина

$$\boldsymbol{x}^* \cdot \boldsymbol{x} = \sum_{k=1}^{n} |x_k|^2$$

принимает наименьшее значение.

Теорема 7. *Система линейных алгебраических уравнений (6) всегда имеет единственное наилучшее приближенное решение, которое определяется по формуле* $\boldsymbol{x}^0 = \boldsymbol{A}^+ \cdot \boldsymbol{b}$.

Доказательство. Для произвольного столбца $\boldsymbol{x}$ положим $\boldsymbol{u} = \boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}^0 = \boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{A}^+ \cdot \boldsymbol{b}$, $\boldsymbol{v} = \boldsymbol{A} \cdot (\boldsymbol{x}^0 - \boldsymbol{x})$. Тогда $\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{u} + \boldsymbol{v}$. Вычислим $|\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}|^2 = (\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x})^* \cdot (\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}) = (\boldsymbol{u} + \boldsymbol{v})^* \cdot (\boldsymbol{u} + \boldsymbol{v}) = \boldsymbol{u} \cdot \boldsymbol{u}^* + \boldsymbol{v}^* \cdot \boldsymbol{v} + \boldsymbol{u}^* \cdot \boldsymbol{v} + \boldsymbol{v}^* \cdot \boldsymbol{u}$. Покажем, что $\boldsymbol{u}^* \cdot \boldsymbol{v} = \boldsymbol{v}^* \cdot \boldsymbol{u} = 0$. Имеем $\boldsymbol{v}^* \cdot \boldsymbol{u} = (\boldsymbol{x}^0 - \boldsymbol{x})^* \cdot \boldsymbol{A}^* \cdot (\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{A}^+ \cdot \boldsymbol{b}) = (\boldsymbol{x}^0 - \boldsymbol{x})^* \cdot (\boldsymbol{A}^* - \boldsymbol{A}^* \cdot \boldsymbol{A} \cdot \boldsymbol{A}^+) \cdot \boldsymbol{b}$. Используя скелетное разложение $\boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{C}$ и формулы (5), получаем $\boldsymbol{A}^* \cdot \boldsymbol{A} \cdot \boldsymbol{A}^+ = \boldsymbol{C}^* \cdot \boldsymbol{B}^* \cdot \boldsymbol{B} \cdot \boldsymbol{C} \cdot \boldsymbol{C}^* \cdot (\boldsymbol{C} \cdot \boldsymbol{C}^*)^{-1} \cdot (\boldsymbol{B}^* \cdot \boldsymbol{B})^{-1} \cdot \boldsymbol{B}^* = \boldsymbol{C}^* \cdot \boldsymbol{B}^* = \boldsymbol{A}^*$, откуда $\boldsymbol{v}^* \cdot \boldsymbol{u} = 0$. Далее, $\boldsymbol{u}^* \cdot \boldsymbol{v} = (\boldsymbol{v}^* \cdot \boldsymbol{u})^* = 0$. Следовательно, $|\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}|^2 = \boldsymbol{u} \cdot \boldsymbol{u}^* + \boldsymbol{v}^* \cdot \boldsymbol{v} = |\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}^0|^2 + |\boldsymbol{A} \cdot (\boldsymbol{x}^0 - \boldsymbol{x})|^2$, т. е.

$$|\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}|^2 = |\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}^0|^2 + |\boldsymbol{A} \cdot (\boldsymbol{x}^0 - \boldsymbol{x})|^2, \tag{7}$$

и потому для любого столбца $\boldsymbol{x}$ имеем $|\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}| \geq |\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}^0|$.

Пусть для столбца $\boldsymbol{x}$ справедливо $|\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}| = |\boldsymbol{b} - \boldsymbol{A} \cdot \boldsymbol{x}^0|$. Положим $\boldsymbol{z} = \boldsymbol{x} - \boldsymbol{x}^0$. Тогда, согласно (7), $\boldsymbol{A} \cdot \boldsymbol{z} = \boldsymbol{A} \cdot (\boldsymbol{x} - \boldsymbol{x}^0) = 0$. Так как $\boldsymbol{x} = \boldsymbol{x}^0 + \boldsymbol{z}$, имеем $|\boldsymbol{x}|^2 = (\boldsymbol{x}^0 + \boldsymbol{z})^* \cdot (\boldsymbol{x}^0 + \boldsymbol{z}) = |\boldsymbol{x}^0|^2 + |\boldsymbol{z}|^2 + (\boldsymbol{x}^0)^* \cdot \boldsymbol{z} + \boldsymbol{z}^* \cdot \boldsymbol{x}^0$. Покажем, что $(\boldsymbol{x}^0)^* \cdot \boldsymbol{z} = 0$. По определению $\boldsymbol{A}^+ = \boldsymbol{A}^* \cdot \boldsymbol{V}$, следовательно, $(\boldsymbol{x}^0)^* \cdot \boldsymbol{z} = (\boldsymbol{A}^+ \cdot \boldsymbol{b})^* \cdot \boldsymbol{z} = (\boldsymbol{A}^* \cdot \boldsymbol{V} \cdot \boldsymbol{b})^* \cdot \boldsymbol{z} = \boldsymbol{b}^* \cdot \boldsymbol{V}^* \cdot \boldsymbol{A} \cdot \boldsymbol{z} = 0$. Таким образом, $(\boldsymbol{x}^0)^* \cdot \boldsymbol{z} = 0$. Но тогда и $\boldsymbol{z}^* \cdot \boldsymbol{x}^0 = ((\boldsymbol{x}^0)^* \cdot \boldsymbol{z})^* = 0$.

Мы видим, что $|\boldsymbol{x}|^2 = |\boldsymbol{x}^0|^2 + |\boldsymbol{z}|^2$, поэтому $|\boldsymbol{x}| \geq |\boldsymbol{x}|^0$ и равенство достигается только при $\boldsymbol{z} = \mathbf{0}$.

Пусть $\boldsymbol{A} = (a_{jk})_{m\times n}$ – матрица. Определим *норму* матрицы $\boldsymbol{A}$ как неотрицательное число

$$\|\boldsymbol{A}\| = \sqrt{\sum_{j=1}^{m}\sum_{k=1}^{n} |a_{jk}|^2}.$$

Очевидно, что

$$\|\boldsymbol{A}\|^2 = \sum_{j=1}^{m} |\boldsymbol{A}(j\bullet)|^2 = \sum_{k=1}^{n} |\boldsymbol{A}(\bullet k)|^2.$$

Рассмотрим матричное уравнение $\boldsymbol{A}\cdot\boldsymbol{X} = \boldsymbol{B}$, где $\boldsymbol{A}$ и $\boldsymbol{B}$ – известные матрицы размеров $m\times n$ и $m\times\ell$, $\boldsymbol{X}$ – неизвестная матрица размеров $n\times\ell$.

Наилучшим приближенным решением матричного уравнения $\boldsymbol{A}\cdot\boldsymbol{X} = \boldsymbol{B}$ называется такая матрица $\boldsymbol{X}^0$, для которой норма $\|\boldsymbol{B} - \boldsymbol{A}\cdot\boldsymbol{X}\|$ достигает наименьшего значения при $\boldsymbol{X} = \boldsymbol{X}^0$ и для всех матриц $\boldsymbol{X}$, у которых норма $\|\boldsymbol{B} - \boldsymbol{A}\cdot\boldsymbol{X}\|$ минимальна, $\|\boldsymbol{X}^0\| \leq \|\boldsymbol{X}\|$.

Теорема 8. *Матричное уравнение $\boldsymbol{A}\cdot\boldsymbol{X} = \boldsymbol{B}$ всегда имеет единственное наилучшее приближенное решение, которое определяется по формуле $\boldsymbol{X}^0 = \boldsymbol{A}^+\cdot\boldsymbol{B}$.*

Доказательство. Из соотношений

$$\|\boldsymbol{B} - \boldsymbol{A}\cdot\boldsymbol{X}\|^2 = \sum_{k=1}^{n} |\boldsymbol{B}(\bullet k) - \boldsymbol{A}\cdot\boldsymbol{X}(\bullet k)|^2$$

и

$$\|\boldsymbol{X}\|^2 = \sum_{k=1}^{n} |\boldsymbol{X}(\bullet k)|^2$$

следует, что k-й столбец искомой матрицы $\boldsymbol{X}^0$ должен быть наилучшим приближенным решением системы линейных уравнений $\boldsymbol{A}\cdot\boldsymbol{X}(\bullet k) = \boldsymbol{B}(\bullet k)$. По теореме 7 $\boldsymbol{X}^0(\bullet k) = \boldsymbol{A}^+\cdot\boldsymbol{B}(\bullet k)$. Так как это равенство справедливо при $k = 1, 2, \ldots, \ell$, получаем требуемое утверждение.

В случае, когда $\boldsymbol{B} = \boldsymbol{E}_m$ – единичная матрица, имеем $\boldsymbol{X}^0 = \boldsymbol{A}^+$. Следовательно, справедлива

Теорема 9. *Псевдообратная матрица $\boldsymbol{A}^+$ является наилучшим приближенным решением (по методу наименьших квадратов) матричного уравнения $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{E}$.*

Это свойство псевдообратной матрицы может быть принято в качестве ее определения.

Задачи для самостоятельного решения

1. Построить ассоциированные матрицы $\breve{\boldsymbol{A}}_p$ при $p = 2, 3, 4$ для следующих матриц (сочетания упорядочиваются лексикографически):

$$\text{а) } \boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix}; \ \text{б) } \boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 & 1 & 2 \\ 3 & 1 & 2 & 2 & 1 \\ 2 & 3 & 1 & 1 & 2 \\ 4 & 5 & 1 & 2 & 3 \\ 5 & 4 & 2 & 3 & 1 \end{pmatrix}.$$

2. Найти псевдообратные матрицы для следующих матриц:

$$\text{а) } \begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 3 & 4 & 5 \\ 3 & 4 & 5 & 6 \\ 4 & 5 & 6 & 7 \end{pmatrix}; \ \text{б) } \begin{pmatrix} 1 & 1 & 1 & 1 \\ 2 & 2 & 2 & 2 \\ 3 & 3 & 3 & 3 \end{pmatrix}; \ \text{в) } \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \end{pmatrix}.$$

3. Найти наилучшее приближенное решение системы линейных алгебраических уравнений:

$$\text{а) } \begin{cases} x_1 + x_2 + x_3 = 1, \\ x_1 + x_2 + x_3 = 2, \\ x_1 + x_2 + x_3 = 3; \end{cases} \quad \text{б) } \begin{cases} x_1 + 2x_2 + 3x_3 + 4x_4 = 10, \\ 2x_1 + 3x_2 + 4x_3 + 5x_4 = 14, \\ 3x_1 + 4x_2 + 5x_3 + 6x_4 = 18. \end{cases}$$

4. Найти наилучшее приближенное решение матричного уравнения $\boldsymbol{A} \cdot \boldsymbol{X} = \boldsymbol{B}$, где

$$\text{а) } \boldsymbol{A} = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 5 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 1 & 2 & 3 \\ 3 & 1 & 2 \\ 2 & 3 & 1 \end{pmatrix};$$

$$\text{б) } \boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 3 \end{pmatrix}, \boldsymbol{B} = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 3 \\ 1 & 2 & 2 \end{pmatrix}.$$

Глава 2

Алгоритм Гаусса и его приложения

§ 2.1. Метод Гаусса

Рассмотрим крамеровскую систему линейных алгебраических уравнений над полем F:

$$\begin{cases} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n = b_1; \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n = b_2; \\ \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \ldots \\ a_{n1}x_1 + a_{n2}x_2 + \ldots + a_{nn}x_n = b_n. \end{cases} \tag{1}$$

Эта система в матричной записи имеет вид

$$\boldsymbol{A} \cdot \boldsymbol{x} = \boldsymbol{b}, \text{ где } \boldsymbol{A} = (a_{tk})_{n\times n},\ \boldsymbol{x} = (x_1, x_2, \ldots, x_n)^\top, \\ \boldsymbol{b} = (b_1, b_2, \ldots, b_n)^\top.$$

Предположим, что $a_{11} \neq 0$, и исключим x_1 из всех уравнений, кроме первого. Затем, предполагая, что во втором уравнении коэффициент при x_2 отличен от 0, исключим x_2 из всех уравнений, начиная с 3-го. Продолжая этот процесс, пока в первом из полученных при очередном шаге уравнений (i-м по порядку) коэффициент при x_i отличен от 0, придем че-

рез p шагов ($p \leq n-1$) к системе с расширенной матрицей
$$\left(\begin{array}{cccccccc|c} a_{11} & a_{12} & \dots & a_{1p} & a_{1,p+1} & \dots & a_{1n} & & b_1 \\ 0 & a_{22}^{(1)} & \dots & a_{2p}^{(1)} & a_{2,p+1}^{(1)} & \dots & a_{2n}^{(1)} & & b_2^{(1)} \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots & & \dots \\ 0 & 0 & \dots & a_{pp}^{(p-1)} & a_{p,p+1}^{(p-1)} & \dots & a_{pn}^{(p-1)} & & b_p^{(p-1)} \\ 0 & 0 & \dots & 0 & a_{p+1,p+1}^{(p)} & \dots & a_{p+1,n}^{(p)} & & b_{p+1}^{(p)} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots & & \dots \\ 0 & 0 & \dots & 0 & a_{n,p+1}^{(p)} & \dots & a_{n,n}^{(p)} & & b_n^{(p)} \end{array}\right).$$

Совершить p шагов в методе Гаусса возможно при выполнении условий

$$a_{11} \neq 0, a_{22}^{(1)} \neq 0, \dots, a_{pp}^{(p-1)} \neq 0. \tag{2}$$

Найдем эквивалентные (2) условия в терминах миноров матрицы $\boldsymbol{A}$. Обозначим расширенную матрицу, которая получается на ℓ-м шаге в методе Гаусса, через $(\boldsymbol{A}_\ell|\boldsymbol{b}_\ell)$, а расширенную матрицу СЛАУ (1) — через $(\boldsymbol{A}|\boldsymbol{b})$. Так как на каждом шаге новая матрица получается путем прибавления i-й строки предыдущей матрицы, умноженной на подходящий скаляр, к каждой ее строке, расположенной ниже i-й, соответствующие миноры этих матриц, расположенные в первых p строках, совпадают. Таким образом, для любых $\ell \leq p$, $1 \leq j_1 < j_2 < \dots < j_\ell \leq n$

$$\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & \ell \\ j_1 & j_2 & \dots & j_\ell \end{pmatrix} = \boldsymbol{A}_p \begin{pmatrix} 1 & 2 & \dots & \ell \\ j_1 & j_2 & \dots & j_\ell \end{pmatrix}. \tag{3}$$

Так как $\boldsymbol{A}_p \begin{pmatrix} 1 & 2 & \dots & \ell \\ 1 & 2 & \dots & \ell \end{pmatrix} = a_{11}a_{22}^{(1)} \dots a_{\ell\ell}^{(\ell-1)}$, условия (2) с учетом (3) равносильны условиям

$$\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & \ell \\ 1 & 2 & \dots & \ell \end{pmatrix} \neq 0,\ \ell = 1, 2, \dots, p. \tag{4}$$

Таким образом,
$\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & \ell \\ 1 & 2 & \dots & \ell \end{pmatrix} = \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & \ell-1 \\ 1 & 2 & \dots & \ell-1 \end{pmatrix} a_{\ell\ell}^{(\ell-1)}$, откуда

$$a_{\ell\ell}^{(\ell-1)} = \frac{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & \ell \\ 1 & 2 & \dots & \ell \end{pmatrix}}{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & \ell-1 \\ 1 & 2 & \dots & \ell-1 \end{pmatrix}} \text{ при } \ell = 2, \dots, p.$$

Так как для $1 \le k_1 < k_2 < \dots < k_p < k_{p+1} \le n$ и $i > p$ справедливо равенство

$$\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & p & i \\ k_1 & k_2 & \dots & k_p & k_{p+1} \end{pmatrix} = \boldsymbol{A}_p\begin{pmatrix} 1 & 2 & \dots & p & i \\ k_1 & k_2 & \dots & k_p & k_{p+1} \end{pmatrix},$$

учитывая вид матрицы $\boldsymbol{A}_p$, имеем

$$\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix} = a_{11}a_{22}^{(1)} \cdots a_{pp}^{(p-1)}a_{ik}^{(p)}.$$

Отсюда, принимая во внимание формулу

$$\boldsymbol{A}_p\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} = a_{11}a_{22}^{(1)} \dots a_{pp}^{(p-1)}, \tag{5}$$

получаем

$$a_{ik}^{(p)} = \frac{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix}}{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}} (i, k = p+1, \dots, n). \tag{6}$$

§ 2.2. Детерминантное тождество Сильвестра

Из равенства $|\boldsymbol{A}| = |\boldsymbol{A}_p|$, с учетом (5), получаем, используя свойство определителя полураспавшейся матрицы,

$$|\boldsymbol{A}| = \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix} \begin{vmatrix} a_{p+1,p+1}^{(p)} & \dots & a_{p+1,n}^{(p)} \\ \dots & \dots & \dots \\ a_{n,p+1}^{(p)} & \dots & a_{nn}^{(p)} \end{vmatrix}. \tag{1}$$

Рассмотрим окаймляющие минор $\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & p \\ 1 & 2 & \dots & p \end{pmatrix}$ определители $b_{ik} = \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & p & i \\ 1 & 2 & \dots & p & k \end{pmatrix}$ $(i, k = p+1, \dots, n)$. Матрицу, составленную из этих определителей, обозначим через

$$\boldsymbol{B} = \begin{pmatrix} b_{p+1,p+1} & \ldots & b_{p+1,n} \\ \ldots & \ldots & \ldots \\ b_{n,p+1} & \ldots & b_{nn} \end{pmatrix}.$$

В силу формул (6) из § 2.1 имеем

$$\begin{vmatrix} a^{(p)}_{p+1,p+1} & \ldots & a^{(p)}_{p+1,n} \\ \ldots & \ldots & \ldots \\ a^{(p)}_{n,p+1} & \ldots & a^{(p)}_{nn} \end{vmatrix} = \frac{\begin{vmatrix} b_{p+1,p+1} & \ldots & b_{p+1,n} \\ \ldots & \ldots & \ldots \\ b_{n,p+1} & \ldots & b_{nn} \end{vmatrix}}{\left[\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix}\right]^{n-p}}. \tag{2}$$

Из формул (1) и (2) получаем $|A| = \dfrac{|\boldsymbol{B}|}{\left[\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix}\right]^{n-p-1}}$,
откуда следует детерминантное тождество Сильвестра

$$|\boldsymbol{B}| = \left[\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix}\right]^{n-p-1} |A|. \tag{3}$$

Это тождество выражает определитель $|\boldsymbol{B}|$, составленный из окаймляющих данный минор определителей, через исходный определитель и окаймляемый минор.

Равенство (3) установлено для матриц $\boldsymbol{A} = (a_{ij})_{n\times n}$, для которых выполнены условия (4) . Покажем, как с помощью соображений непрерывности эти ограничения можно отбросить в случае, когда поле F является полем действительных или комплексных чисел. Пусть $F = \mathbb{R}$ или $F = \mathbb{C}$ и $\boldsymbol{A} = (a_{ij})_{n\times n}$ — матрица, для которой не выполнены условия (4). Зафиксируем число $\varepsilon \in \mathbb{R}$ и положим $\boldsymbol{A}_\varepsilon = \boldsymbol{A} + \varepsilon\boldsymbol{E}$. Тогда $\lim_{\varepsilon\to 0} \boldsymbol{A}_\varepsilon = \boldsymbol{A}$, так как предел матрицы, элементы которой являются функциями, есть матрица, состоящая из пределов этих функций (см. § 7.3).

Минор $\boldsymbol{A}_\varepsilon\begin{pmatrix} 1 & 2 & \ldots & \ell \\ 1 & 2 & \ldots & \ell \end{pmatrix} = \begin{vmatrix} a_{11}+\varepsilon & a_{12} & \ldots & a_{1\ell} \\ a_{21} & a_{22}+\varepsilon & \ldots & a_{2\ell} \\ \ldots & \ldots & \ldots & \ldots \\ a_{\ell 1} & a_{\ell 2} & \ldots & a_{\ell\ell}+\varepsilon \end{vmatrix}$
является многочленом относительно ε со старшим коэффициентом ε^ℓ. При $\ell = 1, 2, \ldots, p$ указанные многочлены имеют конечное множество K комплексных корней. Следовательно, существует

бесконечно малая последовательность $\{\varepsilon_n\}_1^\infty$ такая, что $\varepsilon_n \in \mathbb{R} \setminus K$ при любом натуральном n. Тогда при всех натуральных n и $\ell = 1, 2, \ldots, p$ справедливо $\boldsymbol{A}_{\varepsilon_n}\begin{pmatrix} 1 & 2 & \ldots & \ell \\ 1 & 2 & \ldots & \ell \end{pmatrix} \neq 0$. Поэтому равенство (3) выполняется для каждой матрицы $\boldsymbol{A}_{\varepsilon_n}$. Переходя в обеих частях этого равенства к пределу при $n \longrightarrow \infty$, приходим к требуемому равенству для матрицы $\boldsymbol{A}$.

Приведем детерминантное тождество Сильвестра к виду, удобному для использования. Применяя тождество (3) к определителю $\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & p & i_1 & i_2 & \ldots & i_q \\ 1 & 2 & \ldots & p & j_1 & j_2 & \ldots & j_q \end{pmatrix}$, где $p < i_1 < i_2 < \ldots i_q \leq n$, $p < j_1 < j_2 < \ldots j_q \leq n$, получаем следующий вид детерминантного тождества Сильвестра:

$$\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \ldots & i_q \\ j_1 & j_2 & \ldots & j_q \end{pmatrix} = \left[\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & p \\ 1 & 2 & \ldots & p \end{pmatrix}\right]^{q-1} \times \\ \times \boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & p & i_1 & i_2 & \ldots & i_q \\ 1 & 2 & \ldots & p & j_1 & j_2 & \ldots & j_q \end{pmatrix}. \quad (4)$$

§ 2.3. Разложение в произведение треугольных матриц

Осуществим приведение квадратной матрицы ранга r к ступенчатому виду.

Зафиксируем квадратную матрицу $\boldsymbol{A} = (a_{ij})_{n\times n}$ ранга r. Обозначим $D_k = \boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & k \\ 1 & 2 & \ldots & k \end{pmatrix}$ $(k = 1, 2, \ldots, n)$. Предположим, что $D_k \neq 0$ $(k = 1, 2, \ldots, r)$. Используя элементарные преобразования строк, соответствующие преобразованиям системы линейных уравнений с основной матрицей $\boldsymbol{A}$ при решении ее методом Гаусса, матрицу $\boldsymbol{A}$ можно привести к ступенчатому виду

$$\boldsymbol{G}=\begin{pmatrix} a_{11} & a_{12} & \dots & a_{1r} & a_{1,r+1} & \dots & a_{1n} \\ 0 & a_{22}^{(1)} & \dots & a_{2r}^{(1)} & a_{2,r+1}^{(1)} & \dots & a_{2n}^{(1)} \\ 0 & 0 & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & a_{rr}^{(r-1)} & a_{r,r+1}^{(r-1)} & \dots & a_{rn}^{(r-1)} \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 0 & 0 & \dots & 0 \end{pmatrix}.$$

В силу сказанного выше о возможности совершить p шагов в методе Гаусса и условия $D_k \neq 0$ $(k = 1, 2, \dots, r)$ достаточно использовать лишь преобразования, которые состоят в прибавлении строки с номером i, умноженной на скаляр α, к строке с номером j $(1 \leq i < j \leq n)$. Одно такое преобразование можно получить, умножая слева матрицу $\boldsymbol{A}$ на матрицу $\boldsymbol{E}_n + \alpha \boldsymbol{e}_{ji}$, где $\boldsymbol{e}_{ji}$ – квадратная матрица порядка n, у которой все элементы, кроме элемента в j-й строке и i-м столбце, равны 0, а указанный элемент равен 1. Это нижнетреугольная матрица, имеющая только единицы на главной диагонали.

Последовательность элементарных преобразований строк матрицы $\boldsymbol{A}$, описанную выше, можно осуществить, умножая $\boldsymbol{A}$ слева последовательно на нижнетреугольные матрицы с единицами на главной диагонали. Таким образом, $\boldsymbol{G} = \boldsymbol{W} \cdot \boldsymbol{A}$, где $\boldsymbol{W}$ – произведение упомянутых нижнетреугольных матриц и, следовательно, тоже нижнетреугольная матрица, имеющая только единицы на главной диагонали. Матрицы $\boldsymbol{W}$ и $\boldsymbol{G}$ однозначно определяются по матрице $\boldsymbol{A}$. Матрица $\boldsymbol{W}$ является обратимой, и мы получаем представление матрицы $\boldsymbol{A}$:

$$\boldsymbol{A} = \boldsymbol{W}^{-1} \cdot \boldsymbol{G}, \tag{1}$$

в виде произведения нижнетреугольной матрицы $\boldsymbol{W}^{-1}$ на верхнетреугольную матрицу $\boldsymbol{G}$. Вопрос о разложении матрицы $\boldsymbol{A}$ на множители такого типа полностью решается следующей теоремой.

Теорема 10. *Всякую матрицу* $\boldsymbol{A} = (a_{ij})_{n\times n}$ *ранга* r, *где* $D_k = \boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix} \neq 0$ $(k = 1, 2, \dots, r)$, *можно предста-*

вить в виде произведения:

$$\boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{C} = \begin{pmatrix} b_{11} & 0 & \ldots & 0 \\ b_{21} & b_{22} & \ldots & 0 \\ \ldots & \ldots & \ldots & \ldots \\ b_{n1} & b_{n2} & \ldots & b_{nn} \end{pmatrix} \cdot \begin{pmatrix} c_{11} & c_{12} & \ldots & c_{1n} \\ 0 & c_{22} & \ldots & c_{2n} \\ \ldots & \ldots & \ldots & \ldots \\ 0 & 0 & \ldots & c_{nn} \end{pmatrix}. \qquad (2)$$

При этом

$$b_{11}c_{11} = D_1,\ b_{22}c_{22} = \frac{D_2}{D_1}, \ldots, b_{rr}c_{rr} = \frac{D_r}{D_{r-1}}. \qquad (3)$$

Элементам $b_{11}, \ldots, b_{rr}$, $c_{11}, \ldots, c_{rr}$ можно придать произвольные значения, удовлетворяющие (3). Задание этих элементов однозначно определяет элементы первых r столбцов матрицы $\boldsymbol{B}$ и первых r строк матрицы $\boldsymbol{C}$:

$$b_{\ell k} = b_{kk} \frac{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & k-1 & \ell \\ 1 & 2 & \ldots & k-1 & k \end{pmatrix}}{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & k \\ 1 & 2 & \ldots & k \end{pmatrix}},$$

$$c_{k\ell} = c_{kk} \frac{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & k-1 & k \\ 1 & 2 & \ldots & k-1 & \ell \end{pmatrix}}{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & k \\ 1 & 2 & \ldots & k \end{pmatrix}} \qquad (4)$$

($k = 1, 2, \ldots, r$; $\ell = k+1, \ldots, n$).

Если $r < n$, т. е. $|\boldsymbol{A}| = 0$, то в последних $n - r$ столбцах матрицы $\boldsymbol{B}$ можно все элементы положить равными нулю, а в последних $n - r$ строках матрицы $\boldsymbol{C}$ придать всем элементам произвольные значения, или наоборот, последние $n - r$ строк матрицы $\boldsymbol{C}$ заполнить нулями, а последние $n - r$ столбцов матрицы $\boldsymbol{B}$ взять произвольными.

Доказательство. Возможность представления матрицы $\boldsymbol{A}$ в виде произведения (2) была установлена выше (см. равенство (1)).

Пусть (2) — произвольное представление матрицы $\boldsymbol{A}$ в виде произведения нижнетреугольной матрицы $\boldsymbol{B}$ на верхнетреугольную матрицу $\boldsymbol{C}$. Согласно формуле (5) (см. § 1.1) имеем

$$
\begin{aligned}
&\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & k-1 & \ell \\ 1 & 2 & \ldots & k-1 & k \end{pmatrix} = \\
&\qquad = \sum_{1\le t_1<t_2<\ldots<t_k\le n} \boldsymbol{B}\begin{pmatrix} 1 & 2 & \ldots & k-1 & \ell \\ t_1 & t_2 & \ldots & t_{k-1} & t_k \end{pmatrix} \times \\
&\qquad\qquad \times \boldsymbol{C}\begin{pmatrix} t_1 & t_2 & \ldots & t_{k-1} & t_k \\ 1 & 2 & \ldots & k-1 & k \end{pmatrix}
\end{aligned} \tag{5}
$$

$(k = 1, 2, \ldots, r;\ \ell = k, k+1, \ldots, n)$.

Так как $\boldsymbol{C}$ – верхнетреугольная матрица, в ее первых k столбцах единственный ненулевой минор – это $\boldsymbol{C}\begin{pmatrix} 1 & 2 & \ldots & k \\ 1 & 2 & \ldots & k \end{pmatrix}$. Поэтому (5) превращается в

$$
\begin{aligned}
&\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & k-1 & \ell \\ 1 & 2 & \ldots & k-1 & k \end{pmatrix} = \\
&= \boldsymbol{B}\begin{pmatrix} 1 & 2 & \ldots & k-1 & \ell \\ 1 & 2 & \ldots & k-1 & k \end{pmatrix} \boldsymbol{C}\begin{pmatrix} 1 & 2 & \ldots & k-1 & k \\ 1 & 2 & \ldots & k-1 & k \end{pmatrix} = \\
&\qquad = b_{11}b_{22}\ldots b_{k-1,k-1}b_{\ell k}c_{11}c_{22}\ldots c_{kk}
\end{aligned} \tag{6}
$$

$(k = 1, 2, \ldots, r;\ \ell = k, k+1, \ldots, n)$. Полагая $\ell = k$, получаем при $k = 1, 2, \ldots, r$

$$
D_k = b_{11}b_{22}\ldots b_{kk}c_{11}c_{22}\ldots c_{kk}, \tag{7}
$$

откуда следуют равенства (3).

Если в произведении (2) матрицу $\boldsymbol{B}$ умножить справа на невырожденную диагональную матрицу $\boldsymbol{M}$ с элементами $s_1, \ldots, s_n$ на главной диагонали, а матрицу $\boldsymbol{C}$ умножить слева на матрицу $\boldsymbol{M}^{-1}$, то равенство (2) останется верным. При этом строки матрицы $\boldsymbol{B}$ умножатся на скаляры $s_1, \ldots, s_n$, а столбцы матрицы $\boldsymbol{C}$ умножатся на $s_1^{-1}, \ldots, s_n^{-1}$. Это означает, что элементы $b_{11}, b_{22}, \ldots, b_{rr}$ и $c_{11}, c_{22}, \ldots, c_{rr}$ можно выбирать произвольно с сохранением равенств (3).

Умножив равенство

$$
\boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & k-1 & \ell \\ 1 & 2 & \ldots & k-1 & k \end{pmatrix} = b_{11}b_{22}\ldots b_{k-1,k-1}b_{\ell k}c_{11}c_{22}\ldots c_{kk}
$$

почленно на b_{kk} и используя (7), получаем первое из равенств (4). Второе проверяется аналогично.

Заметим, что при перемножении матриц $\boldsymbol{B}$ и $\boldsymbol{C}$ элементы $b_{k\ell}$ последних $n-r$ столбцов матрицы $\boldsymbol{B}$ и элементы $c_{\ell k}$ последних $n-r$ строк матрицы $\boldsymbol{C}$ перемножаются только между собой. Из равенства (1) следует, что все элементы последних $n-r$ строк матрицы $\boldsymbol{C}$ можно выбрать равными нулю. Тогда элементы последних $n-r$ столбцов матрицы $\boldsymbol{B}$ можно выбрать произвольными. Ясно, что произведение матриц $\boldsymbol{B}$ и $\boldsymbol{C}$ не изменится, если последние $n-r$ столбцов матрицы $\boldsymbol{B}$ взять нулевыми, а элементы последних $n-r$ строк матрицы $\boldsymbol{C}$ произвольными.

Теорема 10 доказана.

Из матричного равенства (2) получается

Следствие 2. *Элементы первых r столбцов матрицы $\boldsymbol{B}$ и первых r строк матрицы $\boldsymbol{C}$ связаны с элементами матрицы $\boldsymbol{A}$ рекуррентными соотношениями:*

$$\begin{aligned} b_{ik} &= \frac{a_{ik} - \sum_{j=1}^{k-1} b_{ij}c_{jk}}{c_{kk}} \quad (i \geq k, i = 1, 2, \ldots, n;\ k = 1, 2, \ldots, r), \qquad (8) \\ c_{ik} &= \frac{a_{ik} - \sum_{j=1}^{k-1} b_{ij}c_{jk}}{b_{ii}} \quad (i \leq k, i = 1, 2, \ldots, r;\ k = 1, 2, \ldots, n). \end{aligned}$$

Соотношениями (8) удобно пользоваться для фактического вычисления элементов матриц $\boldsymbol{B}$ и $\boldsymbol{C}$.

Следствие 3. *Если $\boldsymbol{A}$ – невырожденная матрица порядка n, удовлетворяющая условиям*

$$D_k = \boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & k \\ 1 & 2 & \ldots & k \end{pmatrix} \neq 0 \ \ (k = 1, 2, \ldots, n),$$

то в представлении (2) матрицы $\boldsymbol{B}$ и $\boldsymbol{C}$ определяются однозначно, как только диагональные элементы этих матриц выбраны в соответствии с условиями (3).

Теорема 11. *Всякая квадратная матрица $\boldsymbol{A}$ порядка n и ранга r, у которой $D_k = \boldsymbol{A}\begin{pmatrix} 1 & 2 & \ldots & k \\ 1 & 2 & \ldots & k \end{pmatrix} \neq 0$ $(k = 1, 2, \ldots, r)$, предста-*

вима в виде произведения нижнетреугольной матрицы $\boldsymbol{F}$, диагональной матрицы $\boldsymbol{D} = [D_1, \frac{D_2}{D_1}, \dots, \frac{D_r}{D_{r-1}}, 0, \dots, 0]$ и верхнетреугольной матрицы $\boldsymbol{H}$:

$$\boldsymbol{A} = \boldsymbol{F} \cdot \boldsymbol{D} \cdot \boldsymbol{H} = \tag{9}$$

$$= \begin{pmatrix} 1 & 0 & \dots & 0 \\ f_{21} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ f_{n1} & f_{n2} & \dots & 1 \end{pmatrix} \cdot \boldsymbol{D} \cdot \begin{pmatrix} 1 & h_{12} & \dots & h_{1n} \\ 0 & 1 & \dots & h_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix},$$

где

$$f_{\ell k} = \frac{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & k-1 & \ell \\ 1 & 2 & \dots & k-1 & k \end{pmatrix}}{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}},$$

$$h_{k\ell} = \frac{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & k-1 & k \\ 1 & 2 & \dots & k-1 & \ell \end{pmatrix}}{\boldsymbol{A}\begin{pmatrix} 1 & 2 & \dots & k \\ 1 & 2 & \dots & k \end{pmatrix}} \tag{10}$$

при $k = 1, 2, \dots, r$; $\ell = k+1, \dots, n$ и $f_{\ell k}, h_{k\ell}$ – произвольные при $k = r+1, \dots, n$; $\ell = k+1, \dots, n$.

Доказательство. Применяя к матрице $\boldsymbol{A}$ теорему 10, получаем $\boldsymbol{A} = \boldsymbol{B} \cdot \boldsymbol{C}$, где

$$\boldsymbol{B} = \begin{pmatrix} b_{11} & 0 & \dots & 0 \\ b_{21} & b_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{pmatrix}, \ \boldsymbol{C} = \begin{pmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ 0 & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c_{nn} \end{pmatrix}$$

и при этом $b_{11}c_{11} = D_1$, $b_{22}c_{22} = \frac{D_2}{D_1}, \dots, b_{rr}c_{rr} = \frac{D_r}{D_{r-1}}$. Запишем матрицы $\boldsymbol{B}$ и $\boldsymbol{C}$ в виде

$$\boldsymbol{B} = \begin{pmatrix} 1 & 0 & \dots & 0 \\ f_{21} & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots \\ f_{n1} & f_{n2} & \dots & 1 \end{pmatrix} \cdot \begin{pmatrix} b_{11} & 0 & \dots & 0 \\ 0 & b_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{nn} \end{pmatrix},$$

$$\boldsymbol{C} = \begin{pmatrix} c_{11} & 0 & \dots & 0 \\ 0 & c_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c_{nn} \end{pmatrix} \cdot \begin{pmatrix} 1 & h_{12} & \dots & h_{1n} \\ 0 & 1 & \dots & h_{2n} \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & 1 \end{pmatrix}.$$

Тогда $f_{\ell k}$ и $h_{k\ell}$ будут вычисляться так, как указано в формулировке теоремы 2.

Утверждение теоремы 11 теперь непосредственно получается из того, что матрица $\boldsymbol{D}$ из ее формулировки равна произведению

$$\begin{pmatrix} b_{11} & 0 & \dots & 0 \\ 0 & b_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & b_{nn} \end{pmatrix} \cdot \begin{pmatrix} c_{11} & 0 & \dots & 0 \\ 0 & c_{22} & \dots & 0 \\ \dots & \dots & \dots & \dots \\ 0 & 0 & \dots & c_{nn} \end{pmatrix}.$$

§ 2.4. Блочные матрицы

Часто приходится рассматривать матрицы, разбитые на прямоугольные «клетки» или «блоки». Пусть $\boldsymbol{A} = (a_{ij})_{m\times n}$ – матрица и $m = \sum_{k=1}^{s} m_k$, $n = \sum_{\ell=1}^{t} n_\ell$, где m_k, n_ℓ – натуральные числа. Разобъем матрицу $\boldsymbol{A}$ с помощью s горизонтальных и t вертикальных линий на блоки, которые являются матрицами $\boldsymbol{A}_{k\ell}$ размеров $m_k \times n_\ell$: $\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \dots & \boldsymbol{A}_{1t} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \dots & \boldsymbol{A}_{2t} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \dots & \boldsymbol{A}_{st} \end{pmatrix}$.

Будем говорить, что матрица $\boldsymbol{A}$ представляется в виде блочной матрицы, и писать $\boldsymbol{A} = (\boldsymbol{A}_{k\ell})_{s\times t}$. Естественным образом определяется одинаковое разбиение на блоки для двух матриц совпадающих размеров.

Пусть $\boldsymbol{A} = (\boldsymbol{A}_{k\ell})_{s\times t}$, $\boldsymbol{B} = (\boldsymbol{B}_{k\ell})_{s\times t}$ – матрицы размеров $m \times n$, имеющие одинаковое разбиение на блоки. Сумма блочных матриц $\boldsymbol{A}$ и $\boldsymbol{B}$ через разбиение на блоки определяется естественным способом: $\boldsymbol{A}+\boldsymbol{B} = (\boldsymbol{A}_{k\ell}+\boldsymbol{B}_{k\ell})_{s\times t}$. Аналогично определяется произведение блочной матрицы на число: $t\boldsymbol{A} = (t\boldsymbol{A}_{k\ell})_{s\times t}$.

Для определения произведения двух блочных матриц предположим, что блочные матрицы $\boldsymbol{A} = (\boldsymbol{A}_{k\ell})_{s\times t}$ и $\boldsymbol{B} = (\boldsymbol{B}_{\ell p})_{t\times q}$ разбиты на блоки так, что количество t групп столбцов матрицы $\boldsymbol{A}$ и групп строк $\boldsymbol{B}$ одно и то же и количество столбцов в блоке $\boldsymbol{A}_{k\ell}$ матрицы

$\boldsymbol{A}$ равно количеству строк в соответствующем блоке $\boldsymbol{B}_{\ell p}$ матрицы $\boldsymbol{B}$.

В указанных предположениях $\boldsymbol{A} \cdot \boldsymbol{B} = (\boldsymbol{C}_{kp})_{s\times q}$, где $\boldsymbol{C}_{kp} = \sum_{r=1}^{t} \boldsymbol{A}_{kr} \cdot \boldsymbol{B}_{rp}$ $(k = 1, \ldots, s,\ p = 1, \ldots, q)$.

Пусть $\boldsymbol{A} = (\boldsymbol{A}_{k\ell})_{s\times t}$ – блочная матрица.

Если $s = t$, то говорят, что клетки $\boldsymbol{A}_{\ell\ell}$ $(\ell = 1, 2, \ldots, s)$ образуют *главную квазидиагональ* блочной матрицы $\boldsymbol{A}$.

Матрица $\boldsymbol{A}$ называется *верхней* (соотв. *нижней*) *квазитреугольной*, если $s = t$ и $\boldsymbol{A}_{k\ell} = \boldsymbol{O}$ для всех $1 \le \ell < k \le s$ (соотв. $1 \le k < \ell \le s$), т. е. если все клетки, расположенные ниже (соотв. выше) главной квазидиагонали матрицы $\boldsymbol{A}$ – нулевые матрицы.

Матрица $\boldsymbol{A}$ называется *квазидиагональной*, если $s = t$ и $\boldsymbol{A}_{k\ell} = \boldsymbol{O}$ для всех $k \ne \ell$. Квазидиагональную матрицу $\boldsymbol{A}$ с блоками $\boldsymbol{A}_{11}, \ldots, \boldsymbol{A}_{tt}$ на главной квазидиагонали будем обозначать через $[\boldsymbol{A}_{11}, \ldots, \boldsymbol{A}_{tt}]$.

Очевидно, что блочная матрица является квазидиагональной тогда и только тогда, когда она является верхней квазитреугольной и нижней квазитреугольной.

Из определения произведения блочных матриц непосредственно вытекают следующие утверждения.

Наблюдение 2. *При умножении блочной матрицы слева* (*соотв. справа*) *на квазидиагональную матрицу строки* (*соотв. столбцы*) *этой блочной матрицы умножаются слева* (*соотв. справа*) *на соответствующие диагональные клетки квазидиагональной матрицы.*

Произведение двух верхних (*соотв. нижних*) *квазитреугольных матриц является верхней* (*соотв. нижней*) *квазитреугольной матрицей, при этом диагональные блоки произведения получаются путем перемножения соответствующих диагональных блоков сомножителей.*

Перемножение квадратных блочных матриц одного и того же порядка всегда выполнимо, если они разбиты на одинаковые квадратные схемы блоков и в каждом из сомножителей на диагональных местах стоят квадратные матрицы.

Из теоремы Лапласа получается следующее утверждение.

Предложение 4. *Если* $\boldsymbol{A}$ *– квазитреугольная (в частности, квазидиагональная) матрица с квадратными блоками на главной квазидиагонали, то определитель этой матрицы равен произведению определителей диагональных блоков:*

$$|\boldsymbol{A}| = |\boldsymbol{A}_{11}||\boldsymbol{A}_{22}|\cdots|\boldsymbol{A}_{ss}|. \tag{1}$$

Пусть

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \dots & \boldsymbol{A}_{1t} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \dots & \boldsymbol{A}_{2t} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \dots & \boldsymbol{A}_{st} \end{pmatrix} -$$

блочная матрица размеров $m \times n$ и $m = \sum_{k=1}^{s} m_k$, $n = \sum_{\ell=1}^{t} n_\ell$, где m_k, n_ℓ – натуральные числа. Прибавим к k-й блочной строке матрицы $\boldsymbol{A}$ ее ℓ-ю блочную строку, умноженную слева на прямоугольную матрицу $\boldsymbol{X}$ размеров $m_k \times m_\ell$. Получим матрицу

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \dots & \boldsymbol{A}_{1t} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{A}_{k1} + \boldsymbol{X}\cdot\boldsymbol{A}_{\ell 1} & \boldsymbol{A}_{k2} + \boldsymbol{X}\cdot\boldsymbol{A}_{\ell 2} & \dots & \boldsymbol{A}_{kt} + \boldsymbol{X}\cdot\boldsymbol{A}_{\ell t} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{A}_{\ell 1} & \boldsymbol{A}_{\ell 2} & \dots & \boldsymbol{A}_{\ell t} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \dots & \boldsymbol{A}_{st} \end{pmatrix}.$$

Рассмотрим вспомогательную квадратную матрицу $\boldsymbol{V}$ порядка m, разбитую на блоки следующим образом:

$$\boldsymbol{V} = \begin{pmatrix} \boldsymbol{E}_{m_1} & \dots & \boldsymbol{O} & \dots & \boldsymbol{O} & \dots & \boldsymbol{O} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \boldsymbol{O} & \dots & \boldsymbol{E}_{m_k} & \dots & \boldsymbol{X} & \dots & \boldsymbol{O} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \boldsymbol{O} & \dots & \boldsymbol{O} & \dots & \boldsymbol{E}_{m_\ell} & \dots & \boldsymbol{O} \\ \dots & \dots & \dots & \dots & \dots & \dots & \dots \\ \boldsymbol{O} & \dots & \boldsymbol{O} & \dots & \boldsymbol{O} & \dots & \boldsymbol{E}_{m_s} \end{pmatrix}.$$

Диагональные клетки матрицы $\boldsymbol{V}$ суть единичные матрицы порядков $m_1, m_2, \dots, m_s$ соответственно. Все недиагональные блоки матрицы $\boldsymbol{V}$ нулевые, за исключением блока $\boldsymbol{X}$, стоящего на пересечении k-й блочной строки и ℓ-го блочного столбца. Ясно, что $|\boldsymbol{V}| = 1$.

Легко видеть, что $\boldsymbol{B} = \boldsymbol{V} \cdot \boldsymbol{A}$. Так как $\boldsymbol{V}$ – невырожденная матрица, $\mathsf{r}(\boldsymbol{B}) = \mathsf{r}(\boldsymbol{A})$. Если матрица $\boldsymbol{A}$ является квадратной, то дополнительно $|\boldsymbol{B}| = |\boldsymbol{V}||\boldsymbol{A}| = |\boldsymbol{A}|$.

К подобным выводам можно прийти, прибавляя к блочному столбцу блочной матрицы $\boldsymbol{A}$ другой ее блочный столбец, умноженный справа на матрицу $\boldsymbol{X}$ подходящих размеров. Таким образом, справедлива следущая

Теорема 12. *Если в блочной матрице $\boldsymbol{A}$ к k-й блочной строке (соотв. столбцу) прибавить ℓ-ю блочную строку (соотв. столбец), умноженную слева (соотв. справа) на прямоугольную матрицу $\boldsymbol{X}$ соответствующих размеров, то при этом преобразовании не изменится ранг матрицы $\boldsymbol{A}$, а в случае, когда $\boldsymbol{A}$ – квадратная матрица, также и определитель матрицы $\boldsymbol{A}$.*

Рассмотрим случай, когда в блочной матрице

$$\boldsymbol{A} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \dots & \boldsymbol{A}_{1t} \\ \boldsymbol{A}_{21} & \boldsymbol{A}_{22} & \dots & \boldsymbol{A}_{2t} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{A}_{s1} & \boldsymbol{A}_{s2} & \dots & \boldsymbol{A}_{st} \end{pmatrix}$$

диагональный блок $\boldsymbol{A}_{11}$ является квадратной невырожденной матрицей. Прибавим к k-й блочной строке матрицы $\boldsymbol{A}$ ее первую строку, умноженную слева на $-\boldsymbol{A}_{k1} \cdot \boldsymbol{A}_{11}^{-1}$ $(k = 2, \dots, s)$. Получим матрицу

$$\boldsymbol{B} = \begin{pmatrix} \boldsymbol{A}_{11} & \boldsymbol{A}_{12} & \dots & \boldsymbol{A}_{1t} \\ \boldsymbol{O} & \boldsymbol{A}_{22}^{(1)} & \dots & \boldsymbol{A}_{2t}^{(1)} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{O} & \boldsymbol{A}_{s2}^{(1)} & \dots & \boldsymbol{A}_{st}^{(1)} \end{pmatrix},$$

где $\boldsymbol{A}_{k\ell}^{(1)} = \boldsymbol{A}_{k\ell} - \boldsymbol{A}_{k1} \cdot \boldsymbol{A}_{11}^{-1} \cdot \boldsymbol{A}_{1\ell}$ $(k = 2, \dots, s,\ \ell = 2, \dots, t)$.

Если $\boldsymbol{A}_{22}^{(1)}$ – квадратная невырожденная матрица, то этот процесс можно продолжить. Таким образом, мы приходим к *обобщенному алгоритму Гаусса.*

Пусть $\boldsymbol{A}$ – квадратная матрица. Тогда

$$|\boldsymbol{A}| = |\boldsymbol{A}_{11}| \begin{vmatrix} \boldsymbol{A}_{22}^{(1)} & \dots & \boldsymbol{A}_{2t}^{(1)} \\ \cdots & \cdots & \cdots \\ \boldsymbol{A}_{s2}^{(1)} & \dots & \boldsymbol{A}_{st}^{(1)} \end{vmatrix},$$

т. е. вычисление определителя матрицы, состоящей из st блоков, сводится к вычислению определителей меньшего порядка.

Рассмотрим определитель матрицы, разбитой на 4 блока:

$$\Delta = \begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix},$$

где $\boldsymbol{A}$ и $\boldsymbol{D}$ – квадратные матрицы.

Пусть $|\boldsymbol{A}| \neq 0$. Прибавим к второй строке первую, умноженную слева на $-\boldsymbol{C} \cdot \boldsymbol{A}^{-1}$. Получим

$$\Delta = \begin{vmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{D} - \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B} \end{vmatrix} = |\boldsymbol{A}||\boldsymbol{D} - \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B}|.$$

Аналогично, если $|\boldsymbol{D}| \neq 0$, то прибавим к первой строке вторую, умноженную слева на $-\boldsymbol{B} \cdot \boldsymbol{D}^{-1}$. Получим

$$\Delta = \begin{vmatrix} \boldsymbol{A} - \boldsymbol{B} \cdot \boldsymbol{D}^{-1} \cdot \boldsymbol{C} & \boldsymbol{O} \\ \boldsymbol{C} & \boldsymbol{D} \end{vmatrix} = |\boldsymbol{A} - \boldsymbol{B} \cdot \boldsymbol{D}^{-1} \cdot \boldsymbol{C}||\boldsymbol{D}|.$$

В частном случае, когда все четыре матрицы $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{C}$, $\boldsymbol{D}$ – квадратные одного порядка n, получаем формулы И. Шура, сводящие вычисление определителя порядка $2n$ к определителям порядка n.

$$\Delta = |\boldsymbol{A} \cdot \boldsymbol{D} - \boldsymbol{A} \cdot \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B}| \;\; (|\boldsymbol{A}| \neq 0), \tag{2}$$

$$\Delta = |\boldsymbol{A} \cdot \boldsymbol{D} - \boldsymbol{B} \cdot \boldsymbol{D}^{-1} \cdot \boldsymbol{C} \cdot \boldsymbol{D}| \;\; (|\boldsymbol{D}| \neq 0). \tag{3}$$

Отсюда получаем

$$\Delta = |\boldsymbol{A} \cdot \boldsymbol{D} - \boldsymbol{C} \cdot \boldsymbol{B}| \;\; (|\boldsymbol{A}| \neq 0,\; \boldsymbol{A} \cdot \boldsymbol{C} = \boldsymbol{C} \cdot \boldsymbol{A}), \tag{4}$$

$$\Delta = |\boldsymbol{A} \cdot \boldsymbol{D} - \boldsymbol{B} \cdot \boldsymbol{C}| \;\; (|\boldsymbol{D}| \neq 0,\; \boldsymbol{B} \cdot \boldsymbol{D} = \boldsymbol{D} \cdot \boldsymbol{B}). \tag{5}$$

Используя соображения непрерывности, в случае полей $\mathbb{R}$ или $\mathbb{C}$, ограничения $|\boldsymbol{A}| \neq 0$ в (4) и $|\boldsymbol{D}| \neq 0$ в (5) можно отбросить.

Теорема 13. *Пусть* $\boldsymbol{M} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}$ *– невырожденная матрица порядка* $n + k$*, причем* $\boldsymbol{A}$ *– невырожденная матрица порядка* n*. Тогда матрица* $\boldsymbol{H} = \boldsymbol{D} - \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B}$ *также является обратимой и*

$$\boldsymbol{M}^{-1} = \begin{pmatrix} \boldsymbol{A}^{-1} + \boldsymbol{A}^{-1} \cdot \boldsymbol{B} \cdot \boldsymbol{H}^{-1} \cdot \boldsymbol{C} \cdot \boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1} \cdot \boldsymbol{B} \cdot \boldsymbol{H}^{-1} \\ -\boldsymbol{H}^{-1} \cdot \boldsymbol{C} \cdot \boldsymbol{A}^{-1} & \boldsymbol{H}^{-1} \end{pmatrix}. \quad (6)$$

Формула Фробениуса (6) сводит обращение матрицы порядка $n + k$ к обращению двух матриц порядков n и k и к операциям сложения и умножения матриц размеров $n \times n$, $k \times k$, $n \times k$, $k \times k$.

Доказательство. Применим к матрице $\boldsymbol{M}$ обобщенный алгоритм Гаусса: прибавим к второй блочной строке первую, умноженную слева на $-\boldsymbol{C} \cdot \boldsymbol{A}^{-1}$. Эта операция равносильна умножению матрицы $\boldsymbol{M}$ слева на матрицу $\begin{pmatrix} \boldsymbol{E}_n & \boldsymbol{O} \\ -\boldsymbol{C} \cdot \boldsymbol{A}^{-1} & \boldsymbol{E}_k \end{pmatrix}$. Поэтому

$$\begin{pmatrix} \boldsymbol{E}_n & \boldsymbol{O} \\ -\boldsymbol{C} \cdot \boldsymbol{A}^{-1} & \boldsymbol{E}_k \end{pmatrix} \cdot \boldsymbol{M} = \quad (7)$$

$$= \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{D} - \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B} \end{pmatrix} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{H} \end{pmatrix}.$$

Из равенств (7) следует $|\boldsymbol{M}| = |\boldsymbol{A}||\boldsymbol{H}|$. Так как $|\boldsymbol{M}| \neq 0$ и $|\boldsymbol{A}| \neq 0$, имеем $|\boldsymbol{H}| \neq 0$. Переходя к обратным матрицам в (7), получаем

$$\boldsymbol{M}^{-1} \cdot \begin{pmatrix} \boldsymbol{E}_n & \boldsymbol{O} \\ -\boldsymbol{C} \cdot \boldsymbol{A}^{-1} & \boldsymbol{E}_k \end{pmatrix}^{-1} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{H} \end{pmatrix}^{-1}. \quad (8)$$

Матрицу $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{H} \end{pmatrix}^{-1}$ будем искать в виде $\begin{pmatrix} \boldsymbol{A}^{-1} & \boldsymbol{U} \\ \boldsymbol{O} & \boldsymbol{H}^{-1} \end{pmatrix}$. Из равенства $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{H} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{A}^{-1} & \boldsymbol{U} \\ \boldsymbol{O} & \boldsymbol{H}^{-1} \end{pmatrix} = \begin{pmatrix} \boldsymbol{E}_n & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{E}_k \end{pmatrix}$ находим $\boldsymbol{A} \cdot \boldsymbol{U} + \boldsymbol{B} \cdot \boldsymbol{H}^{-1} = \boldsymbol{O}$, откуда $\boldsymbol{U} = -\boldsymbol{A}^{-1} \cdot \boldsymbol{B} \cdot \boldsymbol{H}^{-1}$. Теперь из равенства (8) вытекает $\boldsymbol{M}^{-1} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{H} \end{pmatrix}^{-1} \cdot \begin{pmatrix} \boldsymbol{E}_n & \boldsymbol{O} \\ -\boldsymbol{C} \cdot \boldsymbol{A}^{-1} & \boldsymbol{E}_k \end{pmatrix} =$
$= \begin{pmatrix} \boldsymbol{A}^{-1} & -\boldsymbol{A}^{-1} \cdot \boldsymbol{B} \cdot \boldsymbol{H}^{-1} \\ \boldsymbol{O} & \boldsymbol{H}^{-1} \end{pmatrix} \cdot \begin{pmatrix} \boldsymbol{E}_n & \boldsymbol{O} \\ -\boldsymbol{C} \cdot \boldsymbol{A}^{-1} & \boldsymbol{E}_k \end{pmatrix}$.

Выполняя умножение в правой части последнего равенства, приходим к формуле Фробениуса (6).

Теорема доказана.

Аналогично теореме 13 можно доказать следующее утверждение.

Теорема 14. *Пусть* $\boldsymbol{M} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}$ *– невырожденная матрица порядка* $n + k$, *причем* $\boldsymbol{D}$ *– невырожденная матрица порядка* k. *Тогда матрица* $\boldsymbol{K} = \boldsymbol{A} - \boldsymbol{B} \cdot \boldsymbol{D}^{-1} \cdot \boldsymbol{C}$ *также является обратимой и*

$$\boldsymbol{M}^{-1} = \begin{pmatrix} \boldsymbol{K}^{-1} & -\boldsymbol{K}^{-1} \cdot \boldsymbol{B} \cdot \boldsymbol{D}^{-1} \\ -\boldsymbol{D}^{-1} \cdot \boldsymbol{C} \cdot \boldsymbol{K}^{-1} & \boldsymbol{D}^{-1} + \boldsymbol{D}^{-1} \cdot \boldsymbol{C} \cdot \boldsymbol{K}^{-1} \cdot \boldsymbol{B} \cdot \boldsymbol{D}^{-1} \end{pmatrix}. \quad (9)$$

Теорема 15. *Если в блочной матрице* $\boldsymbol{R} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{C} & \boldsymbol{D} \end{pmatrix}$ *блок* $\boldsymbol{A}$ *– невырожденная квадратная матрица порядка* n, *то* $\boldsymbol{R}$ *имеет ранг* n *тогда и только тогда, когда* $\boldsymbol{D} = \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B}$.

Доказательство. Применим к матрице $\boldsymbol{R}$ обобщенный алгоритм Гаусса: прибавим к второй блочной строке первую, умноженную слева на $-\boldsymbol{C} \cdot \boldsymbol{A}^{-1}$. Получим матрицу $\boldsymbol{T} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ \boldsymbol{O} & \boldsymbol{D} - \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B} \end{pmatrix}$. Согласно теореме 12, $\mathsf{r}(\boldsymbol{T}) = \mathsf{r}(\boldsymbol{R})$. Так как $\mathsf{r}(\boldsymbol{A}) = n$, при $\boldsymbol{D} - \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B} \neq \boldsymbol{O}$ будем иметь $\mathsf{r}(\boldsymbol{T}) > n$. Следовательно, $\mathsf{r}(\boldsymbol{R}) = n$ тогда и только тогда, когда $\boldsymbol{D} - \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B} = \boldsymbol{O}$. Теорема доказана.

Из теоремы 15 получается такой алгоритм для вычисления произведения $\boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B}$, где $\boldsymbol{A}$ – невырожденная квадратная матрица порядка n, $\boldsymbol{C}$ и $\boldsymbol{B}$ – произвольные матрицы размеров $k \times n$ и $n \times m$. Следует привести блочную матрицу $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ -\boldsymbol{C} & \boldsymbol{O} \end{pmatrix}$ с помощью алгоритма Гаусса (используя элементарные преобразования строк) к виду $\begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{B}_1 \\ \boldsymbol{O} & \boldsymbol{X} \end{pmatrix}$. Тогда $\boldsymbol{X} = \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B}$.

Для обоснования этого алгоритма применим описанное в нем преобразование к блочной матрице $\boldsymbol{R} = \begin{pmatrix} \boldsymbol{A} & \boldsymbol{B} \\ -\boldsymbol{C} & -\boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B} \end{pmatrix}$.

Получим матрицу $\boldsymbol{T} = \begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{B}_1 \\ \boldsymbol{O} & \boldsymbol{X} - \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B} \end{pmatrix}$. Согласно теореме 15 имеем $\mathsf{r}(\boldsymbol{R}) = n$. Тогда по теореме 12 и $\mathsf{r}(\boldsymbol{T}) = n$. Поскольку $\mathsf{r}(\boldsymbol{A}_1) = \mathsf{r}(\boldsymbol{A}) = n$, заключаем, что $\boldsymbol{X} - \boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B} = \boldsymbol{O}$.

На этом основан еще один алгоритм обращения матрицы. Пусть $\boldsymbol{A}$ – невырожденная квадратная матрица порядка n. Следует привести блочную матрицу $\begin{pmatrix} \boldsymbol{A} & \boldsymbol{E}_n \\ -\boldsymbol{E}_n & \boldsymbol{O} \end{pmatrix}$ с помощью алгоритма Гаусса (используя элементарные преобразования строк) к виду $\begin{pmatrix} \boldsymbol{A}_1 & \boldsymbol{B}_1 \\ \boldsymbol{O} & \boldsymbol{X} \end{pmatrix}$. Тогда $\boldsymbol{X} = \boldsymbol{A}^{-1}$.

Обоснование этого алгоритма непосредственно получается из предыдущего алгоритма.

Рассмотрим пример. Найти с помощью алгоритма матрицу $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}^{-1}$.

Имеем

$$\begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 2 & 3 & 4 & 0 & 1 & 0 \\ 3 & 4 & 6 & 0 & 0 & 1 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & -2 & -2 & 1 & 0 \\ 0 & -2 & -3 & -3 & 0 & 1 \\ 0 & 2 & 3 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix} \sim$$

$$\sim \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \\ 0 & 0 & 0 & -2 & 0 & 1 \\ 0 & 0 & 2 & 2 & -1 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix} \sim \begin{pmatrix} 1 & 2 & 3 & 1 & 0 & 0 \\ 0 & 1 & 2 & 2 & -1 & 0 \\ 0 & 0 & 1 & 1 & -2 & 1 \\ 0 & 0 & 0 & -2 & 0 & 1 \\ 0 & 0 & 0 & 0 & 3 & -2 \\ 0 & 0 & 0 & 1 & -2 & 1 \end{pmatrix}.$$

Таким образом, $\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \\ 3 & 4 & 6 \end{pmatrix}^{-1} = \begin{pmatrix} -2 & 0 & 1 \\ 0 & 3 & -2 \\ 1 & -2 & 1 \end{pmatrix}$.

Задачи для самостоятельного решения

1. Разложить в произведение нижнетреугольной и верхетреугольной матриц следующие матрицы:

$$\text{а)} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix}; \text{ б)} \begin{pmatrix} 1 & 2 & 3 & 1 & 2 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 6 & 7 \\ 4 & 5 & 6 & 7 & 8 \\ 5 & 6 & 7 & 8 & 9 \end{pmatrix}.$$

2. Используя формулы И. Шура, вычислить определители:

$$\text{а)} \begin{vmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 \end{vmatrix}; \text{ б)} \begin{vmatrix} 1 & 2 & 3 & 1 & 2 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 2 & 1 \\ 4 & 5 & 2 & 1 & 3 \\ 5 & 6 & 8 & 7 & 9 \end{vmatrix}.$$

3. Используя формулы Фробениуса, найти обратные матрицы для данных матриц:

$$\text{а)} \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 3 & 4 & 5 & 1 \\ 3 & 4 & 5 & 1 & 2 \\ 4 & 5 & 1 & 2 & 3 \\ 5 & 1 & 2 & 3 & 4 \end{pmatrix}; \text{ б)} \begin{pmatrix} 1 & 2 & 3 & 1 & 2 \\ 2 & 3 & 4 & 5 & 6 \\ 3 & 4 & 5 & 2 & 1 \\ 4 & 5 & 2 & 1 & 3 \\ 5 & 6 & 8 & 7 & 9 \end{pmatrix}.$$

4. Используя алгоритм на основе теоремы 15, вычислить произведение $\boldsymbol{C} \cdot \boldsymbol{A}^{-1} \cdot \boldsymbol{B}$, где

а) $\boldsymbol{A} = \begin{pmatrix} 1 & 2 & 1 \\ 5 & 1 & 4 \\ 7 & 4 & 6 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 5 & 7 & 6 & 8 \\ 1 & 1 & 4 & 3 \\ 7 & 9 & 2 & 8 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} 9 & 2 & 11 \\ 8 & 4 & 5 \end{pmatrix}$;

б) $\boldsymbol{A} = \begin{pmatrix} 4 & 2 & 1 \\ 1 & 1 & 5 \\ 7 & 4 & 6 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 1 & 2 & 6 & 8 \\ 1 & 1 & 4 & 3 \\ 5 & 9 & 2 & 1 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} 0 & 2 & 11 \\ 8 & 7 & 5 \end{pmatrix}$.

5. Вычислить обратные для данных матриц, используя алгоритм на основе теоремы 15:

а) $\begin{pmatrix} 1 & 2 & 1 \\ 5 & 1 & 4 \\ 7 & 4 & 6 \end{pmatrix}$; б) $\begin{pmatrix} 4 & 2 & 1 \\ 1 & 1 & 5 \\ 7 & 4 & 6 \end{pmatrix}$.

Глава 3

Многочленные матрицы

§ 3.1. Элементарные преобразования

Пусть F – поле, n – натуральное число. Квадратная матрица порядка n над кольцом многочленов $F[x]$ называется *многочленной матрицей* или *x-матрицей*. Определим следующие элементарные преобразования x-матриц:

1. Умножение какой-либо строки x-матрицы на ненулевой скаляр из поля F.

2. Прибавление к одной строке x-матрицы какой-либо другой ее строки, умноженной на многочлен $f(x) \in F[x]$.

3. Умножение какого-либо столбца x-матрицы на ненулевой скаляр из поля F.

4. Прибавление к одному столбцу x-матрицы какого-либо другого ее столбца, умноженного на многочлен $f(x) \in F[x]$.

Читателю предлагается показать, что элементарными преобразованиями 1, 2 (соотв. 3, 4) можно переставить любые две строки (соотв. два столбца) x-матрицы.

Говорят, что x-матрица $\boldsymbol{A}$ *эквивалентна* x-матрице $\boldsymbol{B}$, если $\boldsymbol{A}$ можно получить из $\boldsymbol{B}$ конечной цепочкой элементарных преобразований. Имеет место следующая

Теорема 16. *Определенное только что отношение является отношением эквивалентности на множестве $F[x]^{n\times n}$ всех x-матриц фиксированного порядка n.*

Доказательство. Легко понять, что рассматриваемое отношение транзитивно. Поскольку каждое элементарное преобразование очевидным образом обратимо, это отношение симметрично. Оно рефлексивно, так как к любой x-матрице можно применить элементарное преобразование – умножить первую строку на скаляр 1.

Таким образом, $F[x]^{n\times n}$ разбивается на классы эквивалентности по этому отношению. Найдем систему представителей для этих классов эквивалентности.

x-матрица называется *канонической диагональной*, если она диагональная и ее диагональные элементы $f_{11}(x), \dots, f_{nn}(x)$, отличные от нуля, имеют старший коэффициент 1 и каждый многочлен $f_{ii}(x)$ является делителем следующего $f_{i+1,i+1}(x)$. Из этого определения следует, что если среди диагональных элементов канонической диагональной матрицы имеются нули, то они должны занимать последние места. Если же среди диагональных элементов встречаются скаляры, отличные от нуля, то они должны быть равны 1 и занимать первые места.

Имеет место следующая

Теорема 17. *Всякая x-матрица может быть приведена к канонической диагональной форме конечной цепочкой элементарных преобразований.*

Доказательство. Пусть $\boldsymbol{A}$ – x-матрица порядка n. Докажем теорему индукцией по n. При $n = 1$ утверждение очевидно, так как матрица первого порядка является канонической диагональной матрицей. Предположим, что утверждение теоремы справедливо для всех x-матриц порядков $1 \le k < n$. Если $\boldsymbol{A}$ – нулевая матрица, то она является канонической диагональной матрицей.

Пусть $\boldsymbol{A}$ – ненулевая x-матрица. В классе эквивалентности, содержащем $\boldsymbol{A}$, выберем x-матрицу

$$\boldsymbol{B} = \begin{pmatrix} b_{11}(x) & b_{12}(x) & \dots & b_{1n}(x) \\ b_{21}(x) & b_{22}(x) & \dots & b_{2n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ b_{n1}(x) & b_{n2}(x) & \dots & b_{nn}(x) \end{pmatrix},$$

у которой элемент $b_{11}(x)$ отличен от нуля и имеет наименьшую возможную степень. Если старший коэффициент α многочлена $b_{11}(x)$

отличен от 1, то умножим первую строку x-матрицы $\boldsymbol{B}$ на α^{-1}. Будем считать, что старший коэффициент многочлена $b_{11}(x)$ равен 1.

Покажем, что $b_{11}(x)$ делит все элементы первой строки и первого столбца x-матрицы $\boldsymbol{B}$. Пусть $b_{1i}(x) = q_i(x)b_{11}(x) + r_i(x)$, где $r_i(x) = 0$ или $r_i(x) \neq 0$ и $\mathsf{deg}(r_i(x)) < \mathsf{deg}(b_{11}(x))$. Во втором случае умножим первый столбец x-матрицы $\boldsymbol{B}$ на $-q_i(x)$ и прибавим к i-му столбцу, а затем поменяем местами первый и i-й столбцы. У полученной матрицы в левом верхнем углу окажется многочлен $r_i(x)$, что противоречит выбору x-матрицы $\boldsymbol{B}$. Следовательно, $r_i(x) = 0$ и все элементы первой строки x-матрицы $\boldsymbol{B}$ делятся на $b_{11}(x)$. Аналогично доказывается, что $b_{11}(x)$ делит все элементы первого столбца x-матрицы $\boldsymbol{B}$.

Умножая первый столбец x-матрицы $\boldsymbol{B}$ на подходящие многочлены и прибавляя к остальным столбцам, а затем проделывая аналогичные преобразования с первой строкой, из x-матрицы $\boldsymbol{B}$ получим эквивалентную ей x-матрицу

$$\boldsymbol{C} = \begin{pmatrix} b_{11}(x) & 0 & \dots & 0 \\ 0 & c_{22}(x) & \dots & c_{2n}(x) \\ \cdots & \cdots & \cdots & \cdots \\ 0 & c_{n2}(x) & \dots & c_{nn}(x) \end{pmatrix}.$$

Убедимся, что $b_{11}(x)$ делит все элементы $c_{ij}(x)$ $(i, j = 2, \dots, n)$. Для этого прибавим i-ю строку x-матрицы $\boldsymbol{C}$ к ее первой строке. Получим матрицу, аналогичную по свойствам x-матрице $\boldsymbol{B}$. Поэтому рассуждения, проведенные выше, показывают, что $b_{11}(x)$ делит все элементы $c_{ij}(x)$ $(i, j = 2, \dots, n)$.

Положим

$$\boldsymbol{C}_1 = \begin{pmatrix} c_{22}(x) & \dots & c_{2n}(x) \\ \cdots & \cdots & \cdots \\ c_{n2}(x) & \dots & c_{nn}(x) \end{pmatrix}.$$

Легко видеть, что каждое элементарное преобразование x-матрицы $\boldsymbol{C}_1$ можно рассматривать и как элементарное преобразование x-матрицы $\boldsymbol{C}$. Применение предположения индукции теперь завершает доказательство теоремы 17.

Из теоремы 17 вытекает

Следствие 4. *Каждый класс эквивалентности x-матриц порядка n содержит по крайней мере одну матрицу в канонической диагональной форме.*

В следующем параграфе будет установлено, что такая матрица определяется однозначно.

Рассмотрим пример. Привести x-матрицу

$$\boldsymbol{A} = \begin{pmatrix} x^2 & x^2 - x & 3x^2 \\ x^2 - x & 3x^2 & x^3 + 4x^2 - 3x \\ x^2 + x & x^2 + x & 3x^2 + 3x \end{pmatrix}$$

к канонической диагональной форме с помощью элементарных преобразований.

Выполняем элементарные преобразования: в матрице $\boldsymbol{A}$ первую строку вычитаем из остальных; в полученной матрице первый столбец умножаем на -2 и -3 и прибавляем к 2-му и 3-му столбцам соответственно; переставляем строки. Далее читателю предлагается объяснить выполняемые преобразования самостоятельно.

$$\boldsymbol{A} \to \begin{pmatrix} x^2 & x^2 - x & 3x^2 \\ -x & 2x^2 + x & x^3 + x^2 - 3x \\ x & 2x & 3x \end{pmatrix} \to$$

$$\to \begin{pmatrix} x^2 & -x^2 - x & 0 \\ -x & 2x^2 + 3x & x^3 + x^2 \\ x & 0 & 0 \end{pmatrix} \to \begin{pmatrix} 0 & -x^2 - x & 0 \\ 0 & 2x^2 + 3x & x^3 + x^2 \\ x & 0 & 0 \end{pmatrix} \to$$

$$\to \begin{pmatrix} x & 0 & 0 \\ 0 & -x^2 - x & 0 \\ 0 & 2x^2 + 3x & x^3 + x^2 \end{pmatrix} \to \begin{pmatrix} x & 0 & 0 \\ 0 & -x^2 - x & 0 \\ 0 & x & x^3 + x^2 \end{pmatrix} \to$$

$$\to \begin{pmatrix} x & 0 & 0 \\ 0 & -x^2 - x & x^2(x+1)^2 \\ 0 & x & 0 \end{pmatrix} \to \begin{pmatrix} x & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & x^2(x+1)^2 \end{pmatrix}.$$

§ 3.2. Делители миноров

Договоримся, что наибольший общий делитель конечного семейства многочленов, не все из которых равны 0, имеет старший коэффициент 1, а если все многочлены нулевые, то их наибольший общий делитель тоже нулевой.

Зафиксируем x-матрицу $\boldsymbol{A}$ порядка n. Через $D_k(\boldsymbol{A})$ обозначим наибольший общий делитель всех миноров порядка k $(1 \leq k \leq n)$ x-матрицы $\boldsymbol{A}$. В частности, $D_1(\boldsymbol{A})$ – наибольший общий делитель всех элементов x-матрицы $\boldsymbol{A}$, а $D_n(\boldsymbol{A}) = \dfrac{|\boldsymbol{A}|}{\alpha}$, где $\alpha = 1$, если $|\boldsymbol{A}| = 0$, и α – старший коэффициент многочлена $|\boldsymbol{A}|$ в противном случае.

Теорема 18. *Если x-матрицы $\boldsymbol{A}$ и $\boldsymbol{B}$ порядка n эквивалентны, то $D_k(\boldsymbol{A}) = D_k(\boldsymbol{B})$ для всех $k = 1, 2. \ldots, n$.*

Доказательство. Ясно, что достаточно доказать утверждение теоремы в случае, когда x-матрицы $\boldsymbol{A}$ получается из $\boldsymbol{B}$ с помощью одного элементарного преобразования. Рассмотрим возможные случаи.

1. Используется элементарное преобразование 1 или 3 (умножение строки или столбца на ненулевой скаляр). При этом соответствующие миноры порядка k x-матриц $\boldsymbol{A}$ и $\boldsymbol{B}$ либо совпадают (если не содержат строки или столбца, умножаемого на скаляр), либо отличаются этим скалярным множителем. Поэтому наибольшие общие делители миноров очевидным образом совпадают.

2. Используется элементарное преобразование 2 или 4 (прибавление к одной строке (столбцу) другой строки (столбца), умноженной на многочлен). Предположим для определенности, что к i-й строке x-матрицы $\boldsymbol{B}$ прибавляется ее ℓ-я строка, умноженная на многочлен $p(x) \in F[x]$, в результате получается x-матрица $\boldsymbol{A}$. Для минора $\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \ldots & i_k \\ j_1 & j_2 & \ldots & j_k \end{pmatrix}$ $(1 \leq i_1 < i_2 < \cdots < i_k \leq n$, $1 \leq j_1 < j_2 < \cdots < j_k \leq n)$ имеются следующие возможности.

1) $i \notin \{i_1, i_2, \ldots, i_k\}$, т. е. i-я строка не входит в рассматриваемый минор. Ясно, что тогда

$$\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \ldots & i_k \\ j_1 & j_2 & \ldots & j_k \end{pmatrix} = \boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \ldots & i_k \\ j_1 & j_2 & \ldots & j_k \end{pmatrix};$$

2) $i, \ell \in \{i_1, i_2, \ldots, i_k\}$, т. е. i-я и ℓ-я строки входят в рассматриваемый минор. Из свойств определителей вытекает, что и в этом случае $\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \ldots & i_k \\ j_1 & j_2 & \ldots & j_k \end{pmatrix} = \boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \ldots & i_k \\ j_1 & j_2 & \ldots & j_k \end{pmatrix}$;

3) $i \in \{i_1, i_2, \ldots, i_k\}$, $\ell \notin \{i_1, i_2, \ldots, i_k\}$, т. е. i-я строка входит ($i = i_m$), а ℓ-я не входит в рассматриваемый минор. По свойствам определителей получаем

$$\boldsymbol{A}\begin{pmatrix} i_1 & i_2 & \ldots & i & \ldots & i_k \\ j_1 & j_2 & \ldots & j_m & \ldots & j_k \end{pmatrix} =$$
$$= \boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \ldots & i & \ldots & i_k \\ j_1 & j_2 & \ldots & j_m & \ldots & j_k \end{pmatrix} \pm$$
$$\pm\, p(x)\boldsymbol{B}\begin{pmatrix} i_1 & i_2 & \ldots & \ell & \ldots & i_k \\ j_1 & j_2 & \ldots & j_m & \ldots & j_k \end{pmatrix}.$$

Отсюда следует, что наибольший общий делитель $D_k(\boldsymbol{B})$, будучи делителем каждого минора порядка k x-матрицы $\boldsymbol{B}$, делит также и каждый минор порядка k x-матрицы $\boldsymbol{A}$. Таким образом, $D_k(\boldsymbol{B})$ делит $D_k(\boldsymbol{A})$. Поскольку рассматриваемое элементарное преобразование обратимо, заключаем, что и $D_k(\boldsymbol{A})$ делит $D_k(\boldsymbol{B})$. Так как старшие коэффициенты этих многочленов равны 1, если многочлены ненулевые, заключаем, что $D_k(\boldsymbol{A}) = D_k(\boldsymbol{B})$.

Теорема 18 доказана.

Пусть $\boldsymbol{A} = [a_1(x), \ldots, a_n(x)]$ – каноническая диагональная x-матрица порядка n, причем для некоторого $r \leq n$ при $i \leq r$ все $a_i(x) \neq 0$, а при $r < n$ и $i > r$ все $a_{ii}(x) = 0$. Так как матрица $\boldsymbol{A}$ диагональная, все ее миноры, отличные от нуля, являются главными. Все миноры порядка больше r очевидным образом равны нулю. Поскольку $a_i(x)$ делит $a_{i+1}(x)$ при $i = 1, 2, \ldots, n-1$, все миноры порядка k x-матрицы $\boldsymbol{A}$ делятся на $\boldsymbol{A}\begin{pmatrix} 1 & 2 & \cdots & k \\ 1 & 2 & \cdots & k \end{pmatrix} = a_1(x)\cdots a_k(x)$. Тогда $D_k(\boldsymbol{A})$ – наибольший общий делитель миноров порядка k – при $k \leq r$ равен $a_1(x)\cdots a_k(x)$, а при $k > r$ равен нулю. Отсюда получаются формулы для диагональных элементов x-матрицы $\boldsymbol{A}$:

$$a_1(x) = D_1(\boldsymbol{A}),\;\; a_i(x) = \frac{D_i(\boldsymbol{A})}{D_{i-1}(\boldsymbol{A})},\;\; i = 2, \ldots, r, \qquad (1)$$
$$a_i(x) = 0,\;\; i > r.$$

Принимая во внимание теоремы 17 и 18, приходим к следующему утверждению – первому критерию эквивалентности x-матриц.

Теорема 19. *Если наибольшие общие делители $D_k(\boldsymbol{B})$ миноров порядка k x-матрицы $\boldsymbol{B}$ отличны от нуля при $k = 1, \ldots, r$ и $D_{r+1}(\boldsymbol{B}) = 0$, то диагональные элементы $a_i(x)$ канонической диагональной x-матрицы $\boldsymbol{A}$, к которой приводится с помощью элементарных преобразований x-матрица $\boldsymbol{B}$, выражаются по формулам (1) и таким образом определяются по x-матрице $\boldsymbol{B}$ однозначно.*

Многочлены $a_1(x), \ldots, a_n(x)$ называются *инвариантными множителями* x-матрицы $\boldsymbol{B}$.

Число r из формулировки теоремы 3 – это ранг x-матрицы $\boldsymbol{B}$.

Теорема 20. *Для того чтобы две x-матрицы порядка n были эквивалентны, необходимо и достаточно, чтобы совпадали их наибольшие общие делители миноров порядка k при $k = 1, 2, \ldots, n$.*

Эту теорему можно сформулировать и в следующем виде.

Теорема 21. *Для того чтобы две x-матрицы порядка n были эквивалентны, необходимо и достаточно, чтобы совпадали их инвариантные множители на k-м месте при $k = 1, 2, \ldots, n$.*

Принципиально иной критерий эквивалентности x-матриц дается следующим утверждением. Заметим, что x-матрица порядка n является обратимой тогда и только тогда, когда ее определитель является ненулевым элементом поля F.

Теорема 22. *Для того чтобы две x-матрицы $\boldsymbol{A}$ и $\boldsymbol{B}$ порядка n были эквивалентны, необходимо и достаточно, чтобы выполнялось равенство $\boldsymbol{B} = \boldsymbol{P} \cdot \boldsymbol{A} \cdot \boldsymbol{Q}$ для некоторых обратимых x-матриц $\boldsymbol{P}$, $\boldsymbol{Q}$.*

Доказательство необходимости. Элементарные преобразования 1 и 3 x-матрицы $\boldsymbol{A}$ порядка n можно реализовать, умножая $\boldsymbol{A}$ слева на подходящие матрицы. Для умножения i-й строки на скаляр α следует умножить $\boldsymbol{A}$ слева на матрицу $\boldsymbol{L}_n(i, \alpha)$, полученную из единичной матрицы $\boldsymbol{E}_n$ заменой i-й единицы на главной диагонали на α. Чтобы прибавить к i-й строке x-матрицы $\boldsymbol{A}$ ее j-ю строку, умноженную на многочлен $f(x)$, следует умножить $\boldsymbol{A}$ слева на матрицу $\boldsymbol{L}_n(i, j, f(x)) = \boldsymbol{E}_n + f(x)\boldsymbol{e}_{ij}$, которая получается из единичной матрицы $\boldsymbol{E}_n$ заменой нуля в i-й строке и j-м столбце на многочлен

$f(x)$. Легко видеть, что $|\boldsymbol{L}_n(i,\alpha)| = \alpha$ и $|\boldsymbol{L}_n(i,j,f(x))| = 1$. Аналогично легко указать матрицы, на которые следует умножить матрицу $\boldsymbol{A}$ справа, чтобы реализовать элементарные преобразования 2 и 4. При этом $\boldsymbol{R}_n(i,\alpha) = \boldsymbol{L}_n(i,\alpha)$ и $\boldsymbol{R}_n(i,j,f(x)) = \boldsymbol{L}_n(j,i,f(x))$.

Теперь докажем необходимость в теореме 22. Если x-матрица $\boldsymbol{A}$ порядка n эквивалентна x-матрице $\boldsymbol{B}$, то $\boldsymbol{B}$ получается из $\boldsymbol{A}$ конечной цепочкой элементарных преобразований. Учитывая сказанное выше, имеем $\boldsymbol{B} = \boldsymbol{P}_1 \cdot \ldots \cdot \boldsymbol{P}_s \cdot \boldsymbol{A} \cdot \boldsymbol{Q}_1 \cdot \ldots \cdot \boldsymbol{Q}_t$, где $\boldsymbol{P}_k$, $\boldsymbol{Q}_m$ – x-матрицы, определители которых есть ненулевые скаляры из поля F. Так как определитель произведения матриц равен произведению определителей и поле не имеет делителей нуля, полагая $\boldsymbol{P} = \boldsymbol{P}_1 \cdot \ldots \cdot \boldsymbol{P}_s$ и $\boldsymbol{Q} = \boldsymbol{Q}_1 \cdot \ldots \cdot \boldsymbol{Q}_t$, получаем требуемое равенство $\boldsymbol{B} = \boldsymbol{P} \cdot \boldsymbol{A} \cdot \boldsymbol{Q}$.

Для доказательства достаточности в теореме 22 потребуется следующая

Лемма 2. *Всякая обратимая x-матрица $\boldsymbol{P}$ порядка n может быть представлена в виде произведения $\boldsymbol{P}_1 \cdot \ldots \cdot \boldsymbol{P}_s$ x-матриц, каждая из которых имеет вид $\boldsymbol{L}_n(i,\alpha)$ или $\boldsymbol{L}_n(i,j,f(x))$ для подходящих $1 \le i,j \le n$, $\alpha \in F \backslash \{0\}$ и $f(x) \in F[x]$.*

Доказательство. Из условия леммы следует, что наибольший общий делитель $D_n(\boldsymbol{P})$ миноров порядка n матрицы $\boldsymbol{P}$ равен 1. Так как $D_n(\boldsymbol{P})$ делится на наибольшие общие делители миноров порядка меньше n, заключаем, что $D_k(\boldsymbol{P}) = 1$ для всех $k = 1, 2, \ldots, n$. Поскольку единичная матрица $\boldsymbol{E}_n$ является канонической диагональной матрицей и имеет такой же набор элементарных делителей, согласно теореме 3 матрицы $\boldsymbol{P}$ и $\boldsymbol{E}_n$ эквивалентны. Следовательно, матрица $\boldsymbol{P}$ получается из матрицы $\boldsymbol{E}_n$ с помощью конечного числа элементарных преобразований. Реализуя их с помощью умножения справа и слева на подходящие матрицы $\boldsymbol{L}_n(i,\alpha)$ или $\boldsymbol{L}_n(i,j,f(x))$, получаем $\boldsymbol{P} = \boldsymbol{P}_1 \cdot \ldots \cdot \boldsymbol{P}_k \cdot \boldsymbol{E}_n \cdot \boldsymbol{P}_{k+1} \cdot \ldots \cdot \boldsymbol{P}_s$, откуда вытекает требуемое утверждение.

Теперь проведем доказательство достаточности в теореме 22. Пусть имеет место равенство $\boldsymbol{B} = \boldsymbol{P} \cdot \boldsymbol{A} \cdot \boldsymbol{Q}$ для некоторых обратимых x-матриц $\boldsymbol{P}$, $\boldsymbol{Q}$. Согласно лемме имеем $\boldsymbol{P} = \boldsymbol{P}_1 \cdot \ldots \cdot \boldsymbol{P}_s$, $\boldsymbol{Q} = \boldsymbol{Q}_1 \cdot \ldots \cdot \boldsymbol{Q}_t$, где $\boldsymbol{P}_k$, $\boldsymbol{Q}_\ell$ – x-матрицы, каждая из которых имеет вид $\boldsymbol{L}_n(i,\alpha)$ или $\boldsymbol{L}_n(i,j,f(x))$ для подходящих $1 \le i,j \le n$, $\alpha \in F \backslash \{0\}$ и $f(x) \in F[x]$. Тогда $\boldsymbol{B} = \boldsymbol{P}_1 \cdot \ldots \cdot \boldsymbol{P}_s \cdot \boldsymbol{A} \cdot \boldsymbol{Q}_1 \cdot \ldots \cdot \boldsymbol{Q}_t$, откуда следует, что x-матрица $\boldsymbol{B}$ получается из $\boldsymbol{A}$ с помощью конечной цепочки элементарных преобразований, реализуемых умножениями слева на матрицы $\boldsymbol{P}_s, \ldots, \boldsymbol{P}_1$ и справа на матрицы $\boldsymbol{Q}_1, \ldots, \boldsymbol{Q}_t$. Теорема 22 доказана.

Рассмотрим пример.

Привести x-матрицу $\boldsymbol{A} = \begin{pmatrix} x & 1 & 0 & 0 \\ 0 & x & 1 & 0 \\ 0 & 0 & x & 1 \\ 5 & 4 & 3 & x+2 \end{pmatrix}$ к канонической диагональной форме с помощью делителей миноров.

Так как $\boldsymbol{A}\begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 4 \end{pmatrix} = 1$, делитель миноров $D_3 = 1$. Следовательно, $D_1 = D_2 = 1$. Раскладывая по первой строке, находим $D_4 = |\boldsymbol{A}| = x^4 + 2x^3 - 3x^2 + 4x - 5$. Таким образом, каноническая диагональная форма x-матрицы $\boldsymbol{A}$ имеет вид $[1, 1, 1, x^4 + 2x^3 - 3x^2 + 4x - 5]$.

§ 3.3. Элементарные делители

Если x-матрица $\boldsymbol{A}$ является квазидиагональной, то зависимость инвариантных множителей $\boldsymbol{A}$ от инвариантных множителей ее диагональных клеток оказывается сложной. Для совокупности элементарных делителей x-матрицы, определяемой ниже, эта зависимость оказывается более простой. Пусть F – произвольное поле, $\boldsymbol{A} \in F[x]^{n\times n}$, $a_1(x), \ldots, a_n(x)$ – ее инвариантные множители. Разложим каждый инвариантный множитель $a_i(x)$, отличный от 0 и от 1, на неприводимые множители над полем F: пусть $a_i(x) = p_{i1}(x)^{s_{i1}} \cdots p_{ik_i}(x)^{s_{ik_i}}$. Будем называть степени $p_{ij}(x)^{s_{ij}}$ *элементарными делителями* многочлена $a_i(x)$.

Элементарным делителем x-матрицы $\boldsymbol{A}$ называется степень неприводимого многочлена $p_{ij}(x)^{s_{ij}}$, где $1 \le j \le k_i$. Говорят, что указанный элементарный делитель *относится* к неприводимому множителю $p_{ij}(x)$. *Системой элементарных делителей* x-матрицы $\boldsymbol{A}$ называется последовательность всех ее элементарных делителей, где каждый делитель повторяется столько раз, в скольких инвариантных множителях он встречается.

Например, если $a_1(x) = (x-1)(x+1)$, $a_2(x) = (x-1)(x+1)^2$, $a_3(x) = (x-1)^2(x+1)^2$, то система элементарных делителей будет $(x-1), (x-1), (x+1), (x+1)^2, (x+1)^2, (x-1)^2$. Обычно одинаковые множители записываются рядом, если есть разные степени, то они идут в порядке возрастания, а порядок следования наборов степеней различных неприводимых многочленов несуществен.

Теорема 23. *Порядок n, ранг r и система элементарных делителей x-матрицы $\boldsymbol{A}$ однозначно определяют ее инвариантные множители и, следовательно, определяют $\boldsymbol{A}$ с точностью до эквивалентности.*

Доказательство. Инвариантный множитель $a_r(x)$ есть произведение наивысших степеней всех элементарных делителей. Вычеркнув их из системы элементарных делителей, аналогично получаем $a_{r-1}(x)$. Продолжая этот процесс, получим все инвариантые множители x-матрицы $\boldsymbol{A}$, отличные от 1.

Лемма 3. *Система элементарных делителей произвольной диагональной x-матрицы есть объединение элементарных делителей ее диагональных элементов.*

Доказательство. Пусть $\boldsymbol{A} = [f_1(x), f_2(x), \ldots, f_n(x)]$ – диагональная x-матрица. Без ограничения общности предположим, что все многочлены $f_j(x)$ ненулевые и имеют старший коэффициент 1. Обозначим через $D_k(x)$ наибольший общий делитель миноров k-го порядка x-матрицы $\boldsymbol{A}$. Так как старшие коэффициенты многочленов $f_1(x), f_2(x), \ldots, f_n(x)$ равны 1, имеем $D_n(x) = f_1(x)f_2(x)\cdots f_n(x)$. С другой стороны, $D_n(x) = a_1(x)a_2(x)\cdots a_n(x)$, где $a_1(x), a_2(x), \ldots, a_n(x)$ — инвариантные множители x-матрицы $\boldsymbol{A}$. Обозначим через $p_1(x), p_2(x), \ldots p_s(x)$ все различные неприводимые множители многочленов $f_1(x), f_2(x), \ldots, f_n(x)$. Каждый элементарный делитель x-матрицы $\boldsymbol{A}$ будет степенью одного из этих многочленов. Зафиксируем один из них: $p(x) = p_i(x)$. Выделим в $f_m(x)$ наивысшую степень многочлена $p(x)$: $f_m(x) = p(x)^{t_m} g_m(x)$, где $g_m(x)$ не делится на $p(x)$. Покажем, что $p(x)^{t_m}$ $(m = 1, 2, \ldots, n)$ есть система элементарных делителей, относящихся к неприводимому многочлену $p(x)$.

Так как система элементарных делителей x-матрицы $\boldsymbol{A}$ не зависит от порядка ее строк и столбцов, без ограничения общности предположим, что ее строки и столбцы расположены так, что

$$t_1 \leq t_2 \leq \ldots \leq t_n. \tag{1}$$

Найдем наивысшую степень $p(x)$, входящую в $D_k(x)$. Среди миноров k-го порядка x-матрицы $\boldsymbol{A}$ отличны от нуля только главные миноры, которые равны $f_{\ell_1}(x)\cdots f_{\ell_k}(x)$, т. е. равны $p(x)^{t_{\ell_1}+\ldots+t_{\ell_k}}g_{\ell_1}(x)\cdots g_{\ell_k}(x)$. Ввиду неравенств (1) наименьшую степень $p(x)$ содержит минор $f_1(x)\cdots f_k(x)$, равный $p(x)^{t_1+\ldots+t_k}g_1(x)\cdots g_k(x)$ Следовательно, $D_k(x)$ содержит $p(x)$ в степени $t_1+\ldots+t_k$. Значит, $D_{k-1}(x)$ содержит $p(x)$ в степени $t_1+\ldots+t_{k-1}$, и потому инвариантный множитель $a_k(x)$, равный отношению $\frac{D_k(x)}{D_{k-1}(x)}$, содержит $p(x)$ точно в степени t_k. Следовательно, элементарные делители x-матрицы $\boldsymbol{A}$, относящиеся к неприводимому многочлену $p(x)$, совпадают с элементарными делителями диагональных элементов $\boldsymbol{A}$, являющимися максимальными степенями $p(x)$. Таким образом, лемма доказана.

Теорема 24. *Система элементарных делителей квазидиагональной x-матрицы равна объединению систем элементарных делителей ее диагональных блоков.*

Доказательство. Пусть $\boldsymbol{A}=[\boldsymbol{A}_1,\ \boldsymbol{A}_2,\ldots,\ \boldsymbol{A}_m]$ – квазидиагональная x-матрица. Элементарные преобразования блоков $\boldsymbol{A}_j$ можно рассматривать и как преобразования всей x-матрицы $\boldsymbol{A}$. Эти преобразования не нарушают квазидиагонального вида x-матрицы $\boldsymbol{A}$ и преобразования, совершаемые над одним из блоков, не изменяют остальные. Следовательно, элементарными преобразованиями x-матрицы $\boldsymbol{A}$ можно привести все блоки к диагональному виду. Применение леммы завершает доказательство теоремы.

Рассмотрим два примера нахождения элементарных делителей.

1. Найти элементарные делители x-матрицы

$$\boldsymbol{A}=\begin{pmatrix} x^2+2 & 2x+1 & x^2+1 \\ x^2+4x+4 & 2x+3 & x^2+4x+3 \\ x^2-4x+3 & 2x-1 & x^2-4x+2 \end{pmatrix}.$$

Приведем матрицу $\boldsymbol{A}$ к канонической диагональной форме:

$$\boldsymbol{A}\to\begin{pmatrix} 1 & 2x+1 & x^2+1 \\ 1 & 2x+3 & x^2+4x+3 \\ 1 & 2x-1 & x^2-4x+2 \end{pmatrix}\to\begin{pmatrix} 1 & 2x+1 & x^2+1 \\ 0 & 2 & 4x+2 \\ 0 & -2 & -4x+1 \end{pmatrix}\to$$

$$\to\begin{pmatrix} 1 & 2x+1 & x^2+1 \\ 0 & 2 & 4x+2 \\ 0 & 0 & 3 \end{pmatrix}\to\begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix}.$$

В этом случае элементарных делителей не существует.

2. Найти элементарные делители x-матрицы

$$\boldsymbol{A} = \begin{pmatrix} x^3+2 & x^3+1 \\ 2x^3-x^2-x+3 & 2x^3-x^2-x+2 \end{pmatrix}.$$

Приведем матрицу $\boldsymbol{A}$ к канонической диагональной форме: $\boldsymbol{A} \to$ $\to \begin{pmatrix} 1 & x^3+1 \\ 1 & 2x^3-x^2-x+2 \end{pmatrix} \to \begin{pmatrix} 1 & x^3+1 \\ 0 & x^3-x^2-x+1 \end{pmatrix} \to$ $\to \begin{pmatrix} 1 & 0 \\ 0 & x^3-x^2-x+1 \end{pmatrix}$. Разложим на неприводимые множители $x^3-x^2-x+1 = (x^2-1)(x-1) = (x-1)^2(x+1)$. Элементарные делители: $(x-1)^2$, $x+1$.

Теперь рассмотрим пример нахождения канонической формы квадратной x-матрицы по ее элементарным делителям, рангу и порядку.

Найдем каноническую форму квадратной x-матрицы по ее элементарным делителям $x+1$, $x+1$, $(x+1)^2$, $x-1$, $x-1$, $(x-1)^2$, рангу $r=4$ и порядку $n=5$.

Инвариантные множители:
$a_4 = (x+1)^2(x-1)^2$, $a_3 = (x+1)(x-1)$, $a_2 = (x+1)(x-1)$, $a_1 = 1$.

Искомая форма имеет вид

$$[1, x^2-1, x^2-1, (x^2-1)^2, 0].$$

Задачи для самостоятельного решения

1. Привести следующие x-матрицы к канонической диагональной форме с помощью элементарных преобразований:

а) $\begin{pmatrix} x-2 & -1 & 0 \\ 0 & x-2 & -1 \\ 0 & 0 & x-2 \end{pmatrix}$; б) $\begin{pmatrix} x(x+1) & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & (x+1)^2 \end{pmatrix}$;

в) $\begin{pmatrix} x+1 & x^2+1 & x^2 \\ 3x-1 & 3x^2-1 & x^2+2x \\ x-1 & x^2-1 & x \end{pmatrix}$.

2. Привести следующие x-матрицы к канонической диагональной форме с помощью делителей миноров:

а) $\begin{pmatrix} x(x+1) & 0 & 0 \\ 0 & x & 0 \\ 0 & 0 & (x+1)^2 \end{pmatrix}$;

б) $\boldsymbol{A} = \begin{pmatrix} x+\alpha & \beta & 1 & 0 \\ -\beta & x+\alpha & 0 & 1 \\ 0 & 0 & x+\alpha & \beta \\ 0 & 0 & -\beta & x+\alpha \end{pmatrix}.$

3. Найти элементарные делители следующих x-матриц:

а) $\begin{pmatrix} x^3-2x^2+2x-1 & x^2-2x+1 \\ 2x^3-2x^2+x-1 & 2x^2-2x \end{pmatrix};$

б) $\begin{pmatrix} 2x^2+3 & x^2+1 & x^6+6x^4+x^2+2 \\ 4x^2+11 & 2x^2+5 & 2x^6+12x^4+2x^2-26 \\ 2x^2+3 & x^2+1 & 2x^6+12x^4+x^2-30 \end{pmatrix}.$

4. Найти каноническую диагональную форму квадратной x-матрицы по ее элементарным делителям, рангу и порядку:

а) $x+2$, $(x+2)^2$, $(x+2)^3$, $x-2$, $(x-2)^3$, $r=n=4$;
б) $x-1$, $(x-1)$, $(x-1)^3$, $x+2$, $(x+2)^2$, $r=4$, $n=5$.

Глава 4

Подобие матриц

§ 4.1. Матричные многочлены

Пусть F – поле. Кольцо матриц $F[x]^{n\times n}$ совпадает с кольцом многочленов $F^{n\times n}[x]$. Поэтому x-матрицы можно рассматривать и как матричные многочлены. Естественным образом определяется степень матричного многочлена.

Матричный многочлен $\boldsymbol{A}(x) = \boldsymbol{A}_0 x^m + \boldsymbol{A}_1 x^{m-1} + \ldots + \boldsymbol{A}_m$ из кольца $F^{n\times n}[x]$ называется *регулярным*, если матрица $\boldsymbol{A}_0$ является обратимой.

Легко видеть, что $\mathsf{deg}(\boldsymbol{A}(x)\boldsymbol{B}(x)) = \mathsf{deg}(\boldsymbol{A}(x)) + \mathsf{deg}(\boldsymbol{B}(x))$, если по крайней мере один из матричных многочленов $\boldsymbol{A}(x)$, $\boldsymbol{B}(x)$ регулярен.

Следующее утверждение параллельно теореме о делимости с остатком для обычных многочленов над полем.

Теорема 25. *Для произвольного матричного многочлена $\boldsymbol{A}(x)$ и регулярного матричного многочлена $\boldsymbol{B}(x)$ существуют матричные многочлены $\boldsymbol{P}(x)$, $\boldsymbol{S}(x)$, $\boldsymbol{Q}(x)$, $\boldsymbol{R}(x)$, удовлетворяющие требованиям:*

1) $\boldsymbol{A}(x) = \boldsymbol{B}(x)\boldsymbol{P}(x) + \boldsymbol{S}(x)$, $\boldsymbol{A}(x) = \boldsymbol{Q}(x)\boldsymbol{B}(x) + \boldsymbol{R}(x)$;

2) либо $\boldsymbol{S}(x) = \boldsymbol{O}$, либо степень $\boldsymbol{S}(x)$ меньше степени $\boldsymbol{B}(x)$; либо $\boldsymbol{R}(x) = \boldsymbol{O}$, либо степень $\boldsymbol{R}(x)$ меньше степени $\boldsymbol{B}(x)$.

Требованиями 1) и 2) указанные многочлены определяются однозначно и называются: $\boldsymbol{P}(x)$, $\boldsymbol{S}(x)$ – левыми частным и остат-

ком, $\boldsymbol{Q}(x)$, $\boldsymbol{R}(x)$ *– правыми частным и остатком от деления* $\boldsymbol{A}(x)$ *на* $\boldsymbol{B}(x)$*.*

Доказательство. Для доказательства существования используем индукцию по $\deg(\boldsymbol{A}(x))$. Пусть $\boldsymbol{A}(x) = \boldsymbol{A}_0x^m + \boldsymbol{A}_1x^{m-1} + \ldots + \boldsymbol{A}_m$, $\boldsymbol{B}(x) = \boldsymbol{B}_0x^n + \boldsymbol{B}_1x^{n-1} + \ldots + \boldsymbol{B}_n$. Так как $\boldsymbol{B}_0$ – обратимая матрица, $n = \deg(\boldsymbol{B}(x)) \geq 0$. Если $n = 0$, то положим $\boldsymbol{P}(x) = \boldsymbol{B}_0^{-1}\boldsymbol{A}(x)$, $\boldsymbol{Q}(x) = \boldsymbol{A}(x)\boldsymbol{B}_0^{-1}$, $\boldsymbol{S}(x) = \boldsymbol{R}(x) = \boldsymbol{O}$. Предположим, что $n > 0$. Если $m < n$, то положим $\boldsymbol{P}(x) = \boldsymbol{Q}(x) = \boldsymbol{O}$ и $\boldsymbol{S}(x) = \boldsymbol{R}(x) = \boldsymbol{A}(x)$. Пусть теперь $m \geq n$. Положим $\boldsymbol{A}_{(\ell)}(x) = \boldsymbol{A}(x) - \boldsymbol{B}(x)\boldsymbol{B}_0^{-1}\boldsymbol{A}_0x^{m-n}$, $\boldsymbol{A}_{(r)}(x) = \boldsymbol{A}(x) - \boldsymbol{A}_0\boldsymbol{B}_0^{-1}\boldsymbol{B}(x)x^{m-n}$. Легко понять, что $\deg(\boldsymbol{A}_{(\ell)}(x)) < m$ и $\deg(\boldsymbol{A}_{(r)}(x)) < m$. По предположению индукции существуют такие многочлены $\boldsymbol{P}_1(x)$,$\boldsymbol{Q}_1(x)$, $\boldsymbol{R}(x)$, $\boldsymbol{S}(x)$, что

$$\boldsymbol{A}_{(\ell)}(x) = \boldsymbol{B}(x)\boldsymbol{P}_1(x) + \boldsymbol{S}(x), \boldsymbol{A}_{(r)}(x) = \boldsymbol{Q}_1(x)\boldsymbol{B}(x) + \boldsymbol{R}(x).$$

Тогда

$$\begin{aligned} \boldsymbol{A}(x) &= \boldsymbol{B}(x)(\boldsymbol{P}_1(x) + \boldsymbol{B}_0^{-1}\boldsymbol{A}_0x^{m-n}) + \boldsymbol{S}(x), \\ \boldsymbol{A}(x) &= (\boldsymbol{P}_1(x) + \boldsymbol{A}_0\boldsymbol{B}_0^{-1}x^{m-n})\boldsymbol{B}(x) + \boldsymbol{R}(x). \end{aligned}$$

Этим существование доказано.

Докажем единственность. Предположим, что

$$\boldsymbol{A}(x) = \boldsymbol{B}(x)\boldsymbol{P}_1(x) + \boldsymbol{S}_1(x), \boldsymbol{A}(x) = \boldsymbol{B}(x)\boldsymbol{P}_2(x) + \boldsymbol{S}_2(x).$$

Тогда $\boldsymbol{B}(x)\boldsymbol{P}_1(x)+\boldsymbol{S}_1(x) = \boldsymbol{B}(x)\boldsymbol{P}_2(x)+\boldsymbol{S}_2(x)$, откуда $\boldsymbol{B}(x)(\boldsymbol{P}_1(x) - \boldsymbol{P}_2(x)) = \boldsymbol{S}_2(x) - \boldsymbol{S}_1(x)$. Если $\boldsymbol{P}_1(x) - \boldsymbol{P}_2(x) \neq \boldsymbol{O}$, то приходим к противоречию: степень левой части последнего равенства больше степени его правой части. Следовательно, $\boldsymbol{P}_1(x) - \boldsymbol{P}_2(x) = \boldsymbol{O}$ и $\boldsymbol{S}_2(x) - \boldsymbol{S}_1(x) = \boldsymbol{O}$, т. е. $\boldsymbol{P}_1(x) = \boldsymbol{P}_2(x)$ и $\boldsymbol{S}_1(x) = \boldsymbol{S}_2(x)$. Аналогично доказывается единственность многочленов $\boldsymbol{Q}(x)$, $\boldsymbol{R}(x)$.

Из теоремы 25 непосредственно получается следующее утверждение.

Следствие 5. *Для любого матричного многочлена* $\boldsymbol{A}(x)$ *из* $F[x]^{n\times n}$ *и любого матричного двучлена* $\boldsymbol{E}_nx - \boldsymbol{B}$*, где* $\boldsymbol{B} \in F^{n\times n}$*, существуют однозначно определенные матричные многочлены* $\boldsymbol{P}(x)$*,* $\boldsymbol{S}(x)$ *и матрицы* $\boldsymbol{S}, \boldsymbol{R} \in F^{n\times n}$ *такие, что*

$$\boldsymbol{A}(x) = (\boldsymbol{E}_nx - \boldsymbol{B})\boldsymbol{P}(x) + \boldsymbol{S} \text{ и } \boldsymbol{A}(x) = \boldsymbol{Q}(x)(\boldsymbol{E}_nx - \boldsymbol{B}) + \boldsymbol{R}.$$

Если $\boldsymbol{A}(x) = \boldsymbol{A}_0x^m + \boldsymbol{A}_1x^{m-1} + \ldots + \boldsymbol{A}_m$, *то*

$$\boldsymbol{S} = \boldsymbol{A}_m + \boldsymbol{B}\boldsymbol{A}_{m-1} + \boldsymbol{B}^2\boldsymbol{A}_{m-2} + \ldots + \boldsymbol{B}^m\boldsymbol{A}_0$$

и

$$\boldsymbol{R} = \boldsymbol{A}_m + \boldsymbol{A}_{m-1}\boldsymbol{B} + \boldsymbol{A}_{m-2}\boldsymbol{B}^2 + \ldots + \boldsymbol{A}_0\boldsymbol{B}^m.$$

Для доказательства последнего утверждения запишем $\boldsymbol{P}(x)$ в виде $\boldsymbol{P}_0x^{m-1} + \ldots + \boldsymbol{P}_{m-1}$, затем в равенстве $\boldsymbol{A}_0x^m + \boldsymbol{A}_1x^{m-1} + \ldots + \boldsymbol{A}_m = (\boldsymbol{E}_nx - \boldsymbol{B})(\boldsymbol{P}_0x^{m-1} + \ldots + \boldsymbol{P}_{m-1}) + \boldsymbol{S}$ раскроем скобки в правой части и приравняем коэффициенты при соответствующих степенях x. Получим $\boldsymbol{A}_0 = \boldsymbol{P}_0$, $\boldsymbol{A}_1 = \boldsymbol{P}_1 - \boldsymbol{B}\boldsymbol{P}_0, \ldots$, $\boldsymbol{A}_{m-1} = \boldsymbol{P}_{m-1} - \boldsymbol{B}\boldsymbol{P}_{m-2}$, $\boldsymbol{A}_m = \boldsymbol{S} - \boldsymbol{B}\boldsymbol{P}_{m-1}$. Выразив $\boldsymbol{S}$ из последнего равенства и двигаясь по цепочке равенств справа налево, получим требуемое выражение для матрицы $\boldsymbol{S}$. Для матрицы $\boldsymbol{R}$ доказательство проводится аналогично.

§ 4.2. Скалярная эквивалентность. Критерий подобия

Многочленные матрицы $\boldsymbol{A}(x)$ и $\boldsymbol{B}$ из $F[x]^{n\times n}$ называются *скалярно эквивалентными*, если существуют такие невырожденные матрицы $\boldsymbol{U}, \boldsymbol{V} \in F^{n\times n}$, что $\boldsymbol{A}(x) = \boldsymbol{U}\boldsymbol{B}\boldsymbol{V}$.

Теорема 26. *Два регулярных двучлена эквивалентны тогда и только тогда, когда они скалярно эквивалентны.*

Доказательство. Пусть $\boldsymbol{A}x + \boldsymbol{B}$ и $\boldsymbol{C}x + \boldsymbol{D}$ – регулярные двучлены из $F[x]^{n\times n}$. Они эквивалентны согласно второму критерию эквивалентности x-матриц тогда и только тогда, когда

$$\boldsymbol{A}x + \boldsymbol{B} = \boldsymbol{U}(x)(\boldsymbol{C}x + \boldsymbol{D})\boldsymbol{V}(x) \tag{1}$$

для некоторых обратимых x-матриц $\boldsymbol{U}(x), \boldsymbol{V}(x) \in F(x)^{n\times n}$. Поэтому из скалярной эквивалентности x-матриц следует их эквивалентность. Предположим, что $\boldsymbol{A}x + \boldsymbol{B}$ и $\boldsymbol{C}x + \boldsymbol{D}$ эквивалентны. Тогда выполняется (1) и $|\boldsymbol{U}(x)|, |\boldsymbol{V}(x)| \in F \setminus \{0\}$. Разделим $\boldsymbol{U}(x), \boldsymbol{V}(x)$ на $\boldsymbol{A}x + \boldsymbol{B}$: $\boldsymbol{U}(x) = (\boldsymbol{A}x + \boldsymbol{B})\boldsymbol{P}(x) + \boldsymbol{S}$, $\boldsymbol{V}(x) = \boldsymbol{Q}(x)(\boldsymbol{A}x + \boldsymbol{B}) + \boldsymbol{R}$. Покажем, что матрицы $\boldsymbol{S}, \boldsymbol{R} \in F^{n\times n}$ невырожденные и

$$\boldsymbol{A}x + \boldsymbol{B} = \boldsymbol{S}(\boldsymbol{C}x + \boldsymbol{D})\boldsymbol{R}. \tag{2}$$

Из (1) следует $\boldsymbol{U}(x)^{-1}(\boldsymbol{A}x+\boldsymbol{B}) = (\boldsymbol{C}x+\boldsymbol{D})\boldsymbol{V}(x) = (\boldsymbol{C}x+\boldsymbol{D})\cdot$ $\cdot(\boldsymbol{Q}(x)(\boldsymbol{A}x+\boldsymbol{B})+\boldsymbol{R})$, откуда

$$\boldsymbol{U}(x)^{-1}(\boldsymbol{A}x+\boldsymbol{B}) = (\boldsymbol{C}x+\boldsymbol{D})\boldsymbol{Q}(x)(\boldsymbol{A}x+\boldsymbol{B}) + (\boldsymbol{C}x+\boldsymbol{D})\boldsymbol{R}$$

и

$$(\boldsymbol{U}(x)^{-1} - (\boldsymbol{C}x+D)\boldsymbol{Q}(x))(\boldsymbol{A}x+\boldsymbol{B}) = (\boldsymbol{C}x+\boldsymbol{D})\boldsymbol{R}. \tag{3}$$

Положим $\boldsymbol{T} = \boldsymbol{U}(x)^{-1} - (\boldsymbol{C}x+\boldsymbol{D})\boldsymbol{Q}(x)$. Из равенства (3) следует, что $\boldsymbol{T} \in F^{n\times n}$.

Покажем, что $\boldsymbol{ST} = \boldsymbol{E}_n$. Так как $\boldsymbol{T}(\boldsymbol{A}x+\boldsymbol{B}) = (\boldsymbol{C}x+\boldsymbol{D})\boldsymbol{R}$, тогда (2) будет доказано. Имеем $\boldsymbol{U}(x)^{-1} = \boldsymbol{T} + (\boldsymbol{C}x+\boldsymbol{D})\boldsymbol{Q}(x)$ и далее $\boldsymbol{E}_n = \boldsymbol{U}(x)\boldsymbol{U}(x)^{-1} = \boldsymbol{U}(x)(\boldsymbol{T}+(\boldsymbol{C}x+\boldsymbol{D})\boldsymbol{Q}(x)) = \boldsymbol{U}(x)\boldsymbol{T}+$ $+\boldsymbol{U}(x)(\boldsymbol{C}x+\boldsymbol{D})\boldsymbol{Q}(x)$.

Из (1) получаем $\boldsymbol{U}(x)(\boldsymbol{C}x+\boldsymbol{D}) = (\boldsymbol{A}x+\boldsymbol{B})\boldsymbol{V}(x)^{-1}$. Следовательно, $\boldsymbol{U}(x)\boldsymbol{T}+\boldsymbol{U}(x)(\boldsymbol{C}x+\boldsymbol{D})\boldsymbol{Q}(x) = \boldsymbol{U}(x)\boldsymbol{T}+(\boldsymbol{A}x+\boldsymbol{B})\boldsymbol{V}(x)^{-1}\boldsymbol{Q}(x) =$ $= ((\boldsymbol{A}x+\boldsymbol{B})\boldsymbol{P}(x)+\boldsymbol{S})\boldsymbol{T}+(\boldsymbol{A}x+\boldsymbol{B})\boldsymbol{V}(x)^{-1}\boldsymbol{Q}(x) = (\boldsymbol{A}x+\boldsymbol{B})(\boldsymbol{P}(x)\boldsymbol{T}+$ $+\boldsymbol{V}(x)^{-1}\boldsymbol{Q}(x)) + \boldsymbol{ST}$. Таким образом,

$$\boldsymbol{E}_n = (\boldsymbol{A}x+\boldsymbol{B})(\boldsymbol{P}(x)\boldsymbol{T}+\boldsymbol{V}(x)^{-1}\boldsymbol{Q}(x)) + \boldsymbol{ST},$$

откуда $\boldsymbol{P}(x)\boldsymbol{T}+\boldsymbol{V}(x)^{-1}\boldsymbol{Q}(x) = \boldsymbol{O}$ и $\boldsymbol{ST} = \boldsymbol{E}_n$, что и требуется доказать.

Напомним, что матрицы $\boldsymbol{A}$ и $\boldsymbol{B}$ из $F^{n\times n}$ называются *подобными*, если существует невырожденная матрица $\boldsymbol{T} \in F^{n\times n}$ такая, что $\boldsymbol{A} =$ $= \boldsymbol{T}^{-1}\boldsymbol{BT}$.

Характеристической матрицей матрицы $\boldsymbol{A}$ из $F^{n\times n}$ называется x-матрица $\boldsymbol{E}_n x - \boldsymbol{A}$.

Следующее утверждение дает критерий подобия двух квадратных матриц.

Теорема 27. *Матрицы $\boldsymbol{A}$ и $\boldsymbol{B}$ из $F^{n\times n}$ подобны тогда и только тогда, когда их характеристические матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$ и $\boldsymbol{E}_n x - \boldsymbol{B}$ эквивалентны.*

Доказательство. Если $\boldsymbol{A}$ и $\boldsymbol{B}$ подобны, то $\boldsymbol{A} = \boldsymbol{T}^{-1}\boldsymbol{BT}$ и поэтому $\boldsymbol{E}_n x - \boldsymbol{A} = \boldsymbol{T}^{-1}(\boldsymbol{E}_n x - \boldsymbol{B})\boldsymbol{T}$, т. е. x-матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$ и $\boldsymbol{E}_n x - \boldsymbol{B}$ даже скалярно эквивалентны и поэтому эквивалентны.

Обратно, если x-матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$ и $\boldsymbol{E}_n x - \boldsymbol{B}$ эквивалентны, то в силу теоремы 26 они скалярно эквивалентны и потому $\boldsymbol{E}_n x-$

$-\boldsymbol{A} = \boldsymbol{S}(\boldsymbol{E}_n x - \boldsymbol{B})\boldsymbol{R}$ для некоторых матриц $\boldsymbol{S}, \boldsymbol{R} \in F^{n\times n}$. Тогда $\boldsymbol{E}_n x - \boldsymbol{A} = \boldsymbol{SR}x - \boldsymbol{SBR}$ и $\boldsymbol{SR} = \boldsymbol{E}_n$, $\boldsymbol{A} = \boldsymbol{SBR}$, что и требуется доказать.

Говорят, что матрица $\boldsymbol{T} \in F^{n\times n}$ *осуществляет подобие* матрицы $\boldsymbol{A} \in F^{n\times n}$ матрице $\boldsymbol{B} \in F^{n\times n}$, если $\boldsymbol{A} = \boldsymbol{T}^{-1}\boldsymbol{BT}$.

Чтобы найти такую матрицу для подобных матриц $\boldsymbol{A}, \boldsymbol{B} \in F^{n\times n}$, можно решать систему линейных уравнений относительно n^2 неизвестных элементов матрицы $\boldsymbol{T}$, получаемую из матричного равенства $\boldsymbol{TA} = \boldsymbol{BT}$.

Однако при достаточно больших значениях n это становится слишком трудоемким. Другой способ нахождения матрицы, осуществляющей подобие, состоит в следующем. Находим x-матрицы $\boldsymbol{U}(x), \boldsymbol{V}(x) \in F[x]^{n\times n}$ такие, что $\boldsymbol{E}_n x - \boldsymbol{A} = \boldsymbol{U}(x)(\boldsymbol{E}_n x - \boldsymbol{B})\boldsymbol{V}(x)$. Следуя доказательству теоремы 26, заключаем, что подобие осуществляется, в частности, матрицей $\boldsymbol{R}$ – остатком от деления $\boldsymbol{V}(x)$ справа на $\boldsymbol{E}_n x - \boldsymbol{A}$. Если $\boldsymbol{V}(x) = \boldsymbol{V}_0 + \boldsymbol{V}_1 x + \ldots + \boldsymbol{V}_m x^m$, то

$$\boldsymbol{R} = \boldsymbol{V}_0 + \boldsymbol{V}_1\boldsymbol{A} + \ldots + \boldsymbol{V}_m\boldsymbol{A}^m.$$

Рассмотрим пример. Убедимся, что матрицы $\boldsymbol{A} = \begin{pmatrix} 5 & -1 \\ 9 & -1 \end{pmatrix}$ и $\boldsymbol{B} = \begin{pmatrix} 38 & -81 \\ 16 & -34 \end{pmatrix}$ подобны, и найдем трансформирующую матрицу. Обозначим трансформирующую матрицу через $\boldsymbol{T} = \begin{pmatrix} x & y \\ z & t \end{pmatrix}$. Из равенства $\boldsymbol{AT} = \boldsymbol{TB}$ следует

$$\begin{pmatrix} 5x - z & 5y - t \\ 9x - z & 9y - t \end{pmatrix} = \begin{pmatrix} 38x + 16y & -81x - 34y \\ 38z + 16t & -81z - 34t \end{pmatrix},$$

откуда $\begin{cases} 33x + 16y + z = 0, \\ 81x + 39y - t = 0, \\ 9x - 39z - 16t = 0, \\ 9y + 81z + 33t = 0. \end{cases}$

Решая систему, получаем $\begin{cases} z = -33x - 16y, \\ t = 81x + 33y. \end{cases}$

Следовательно, общий вид трансформирующей матрицы $\boldsymbol{T} = \begin{pmatrix} x & y \\ -33x - 16y & 81x + 33y \end{pmatrix}$ (x, y одновременно не обращаются в нуль).

§ 4.3. Нормальные формы

Пусть $f(x) \in F[x]$, $f(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_{n-1}x^{n-1} + x^n$. *Сопровождающей матрицей* многочлена $f(x)$ называется матрица

$$\boldsymbol{A} = \begin{pmatrix} 0 & 1 & 0 & \ldots & 0 \\ 0 & 0 & 1 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \ldots & -\alpha_{n-1} \end{pmatrix}. \qquad (1)$$

Запишем характеристическую матрицу для матрицы $\boldsymbol{A}$:

$$\boldsymbol{E}_n x - \boldsymbol{A} = \begin{pmatrix} x & -1 & 0 & \ldots & 0 \\ 0 & x & -1 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & -1 \\ \alpha_0 & \alpha_1 & \alpha_2 & \ldots & x + \alpha_{n-1} \end{pmatrix}.$$

Найдем делители миноров этой матрицы. Так как

$$(\boldsymbol{E}_n x - \boldsymbol{A}) \begin{pmatrix} 1 & 2 & \ldots & n-1 \\ 2 & 3 & \ldots & n \end{pmatrix} = (-1)^{n-1},$$

имеем $D_j = 1$ при $j = 1, \ldots, n-1$. Раскладывая определитель

$$\begin{vmatrix} x & -1 & 0 & \ldots & 0 \\ 0 & x & -1 & \ldots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & -1 \\ \alpha_0 & \alpha_1 & \alpha_2 & \ldots & x + \alpha_{n-1} \end{vmatrix}$$

по последней строке, получаем $|\boldsymbol{E}_n x - \boldsymbol{A}| = \alpha_0(-1)^{n+1}(-1)^{n-1} + \alpha_1(-1)^{n+2}(-1)^{n-2}x + \ldots + \alpha_{n-2}(-1)^{2n-1}(-1)x^{n-2} + (x + \alpha_{n-1})x^{n-1} = f(x)$. Следовательно, $D_n = f(x)$. Поэтому инвариантные множители матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$ суть $1, \ldots, 1, f(x)$.

Следовательно, матрица $\boldsymbol{E}_n x - \boldsymbol{A}$ приводится с помощью элементарных преобразований к диагональному виду $[1, \ldots, 1, f(x)]$.

Говорят, что матрица $\boldsymbol{A}$ имеет *естественную нормальную форму*, если она квазидиагональная и ее диагональные клетки $\boldsymbol{A}_1, \boldsymbol{A}_2, \ldots, \boldsymbol{A}_s$ являются сопровождающими матрицами многочленов $f_1(x), f_2(x), \ldots, f_s(x)$ таких, что $f_i(x)$ делит $f_{i+1}(x)$ при $i = 1, \ldots, s-1$.

Пусть n — порядок матрицы $\boldsymbol{A}$, n_i — порядок матрицы $\boldsymbol{A}_i$ при $i = 1, \ldots, s$. Инвариантные множители матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$ суть $1, \ldots, 1, f_1(x), \ldots, f_s(x)$, где количество единиц равно $n-s$. Так как

$$\boldsymbol{E}_n x - \boldsymbol{A} = \begin{pmatrix} \boldsymbol{E}_{n_1} x - \boldsymbol{A}_1 & \boldsymbol{O} & \ldots & \boldsymbol{O} \\ \boldsymbol{O} & \boldsymbol{E}_{n_2} x - \boldsymbol{A}_2 & \ldots & \boldsymbol{O} \\ \ldots & \ldots & \ldots & \ldots \\ \boldsymbol{O} & \boldsymbol{O} & \ldots & \boldsymbol{E}_{n_s} x - \boldsymbol{A}_s \end{pmatrix},$$

приводим каждый блок на главной диагонали с помощью элементарных преобразований к виду $[1, \ldots, 1, f_i(x)]$. Эти преобразования можно выполнять как преобразования всей матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$. Затем, переставляя строки и столбцы, располагаем многочлены в конце главной диагонали по порядку.

Теорема 28. *Каждая матрица $\boldsymbol{A} \in F^{n \times n}$ подобна единственной матрице в естественной нормальной форме.*

Доказательство. Пусть $1, \ldots, 1, f_1(x), \ldots, f_s(x)$ – инвариантные множители матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$, где число единиц равно $n - s$. Тогда $f_1 = D_{n-s+1}$, $f_2 = \frac{D_{n-s+2}}{D_{n-s+1}}, \ldots, f_s = \frac{D_n}{D_{n-1}}$, поэтому $f_1 \ldots f_s = D_n$ и $n = \sum_{i=1}^{s} \mathsf{deg}(f_i)$. С помощью элементарных преобразований матрица $\boldsymbol{E}_n x - \boldsymbol{A}$ приводится к диагональному виду $[1, \ldots, 1, f_1(x), \ldots, f_s(x)]$. Для каждого многочлена f_i построим сопровождающую матрицу $\boldsymbol{B}_i$. Затем построим квазидиагональную матрицу $\boldsymbol{B} = [\boldsymbol{B}_1, \ldots, \boldsymbol{B}_s]$. Она имеет естественную нормальную форму. Так как x-матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$ и $\boldsymbol{E}_n x - \boldsymbol{B}$ эквивалентны, по теореме 27 матрица $\boldsymbol{A}$ подобна матрице $\boldsymbol{B}$.

Единственность следует из того, что естественная нормальная форма $\boldsymbol{B}$ однозначно определяется инвариантными множителями x-матрицы $\boldsymbol{E}_n x – \boldsymbol{A}$, которые однозначно определяются матрицей $\boldsymbol{A}$.

Естественная нормальная форма квазидиагональной матрицы $[\boldsymbol{A}_1 \ldots \boldsymbol{A}_n]$ не получается как $[\boldsymbol{B}_1 \ldots \boldsymbol{B}_n]$ — квазидиагональная матрица из естественных нормальных форм $\boldsymbol{B}_i$ диагональных блоков $\boldsymbol{A}_i$, что является недостатком естественной нормальной формы. Этого недостатка лишена вторая естественная нормальная форма.

Пусть $\boldsymbol{A} \in F^{n\times n}$, $e_1, \ldots, e_s$ – все элементарные делители x-матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$, $\boldsymbol{B}_i$ – сопровождающая матрица многочлена e_i, $i = 1, \ldots, s$. Квазидиагональная матрица $[\boldsymbol{B}_1 \ldots \boldsymbol{B}_s]$ называется *второй естественной нормальной формой* матрицы $\boldsymbol{A}$.

Теорема 29. *Каждая матрица $\boldsymbol{A} \in F^{n\times n}$ подобна матрице $\boldsymbol{B}$ во второй естественной нормальной форме, причем матрица $\boldsymbol{B}$ определена однозначно с точностью до перестановки блоков на главной диагонали.*

Доказательство получается с помощью критерия подобия и теоремы 23 из того факта, что x-матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$ и $\boldsymbol{E}_n x - \boldsymbol{B}$ имеют совпадающие наборы элементарных делителей.

Если характеристический многочлен матрицы $\boldsymbol{A}$ разлагается на линейные множители, то вместо второй естественной нормальной формы используется *жорданова нормальная форма*. Если поле F алгебраически замкнуто, то это справедливо для любой матрицы. В рассматриваемом случае любой элементарный делитель x-матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$ имеет вид $(x - \lambda)^k$. Вместо сопровождающей матрицы для такого многочлена используется *жорданова клетка* порядка k, относящаяся к λ:

$$\boldsymbol{J}_k(\lambda) = \begin{pmatrix} \lambda & 1 & 0 & \ldots & 0 & 0 \\ 0 & \lambda & 1 & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & \lambda & 1 \\ 0 & 0 & 0 & \ldots & 0 & \lambda \end{pmatrix}.$$

Для x-матрицы

$$\boldsymbol{E}_k x - \boldsymbol{J}_k(\lambda) = \begin{pmatrix} x-\lambda & -1 & 0 & \ldots & 0 & 0 \\ 0 & x-\lambda & -1 & \ldots & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \ldots & x-\lambda & -1 \\ 0 & 0 & 0 & \ldots & 0 & x-\lambda \end{pmatrix}$$

имеем $D_{k-1} = 1$, так как $(\boldsymbol{E}_k x - \boldsymbol{J}_k(\lambda)) \begin{pmatrix} 1 & 2 & \ldots & k-1 \\ 2 & 3 & \ldots & k \end{pmatrix} =$ $= (-1)^{k-1}$, и $D_k = (x - \lambda)^k$. Поэтому единственный элементарный делитель x-матрицы $\boldsymbol{E}_k x - \boldsymbol{J}_k(\lambda)$ есть $(x - \lambda)^k$.

Жордановой нормальной формой называется квазидиагональная матрица, на главной диагонали которой расположены жордановы клетки любых порядков, относящиеся к произвольным элементам поля.

Теорема 30. *Если характеристический многочлен матрицы $\boldsymbol{A}$ порядка n разлагается на линейные множители, то она подобна матрице $\boldsymbol{B}$, являющейся жордановой нормальной формой, причем матрица $\boldsymbol{B}$ определена однозначно с точностью до перестановки жордановых клеток на главной диагонали.*

Доказательство. Если характеристический многочлен матрицы $\boldsymbol{A}$ порядка n разлагается на линейные множители, то любой элементарный делитель x-матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$ имеет вид $(x-\lambda)^k$. Построив квазидиагональную матрицу из клеток Жордана, соответствующих всем элементарным делителям x-матрицы $\boldsymbol{E}_n x - \boldsymbol{A}$, получим матрицу $\boldsymbol{B}$. Утверждения о подобии и единственности получаются аналогично теореме 29.

Следствие 6. *Любая матрица $\boldsymbol{A}$ порядка n над алгебраически замкнутым полем F подобна матрице $\boldsymbol{B}$, являющейся жордановой нормальной формой, причем матрица $\boldsymbol{B}$ определена однозначно с точностью до перестановки жордановых клеток на главной диагонали.*

Так как любой многочлен положительной степени над алгебраически замкнутым полем разлагается на линейные множители, остается применить теорему 30.

Поскольку поле комплексных чисел алгебраически замкнуто, сразу получаем

Следствие 7. *Любая матрица $\boldsymbol{A}$ порядка n над полем комплексных чисел подобна матрице $\boldsymbol{B}$, являющейся жордановой нормальной формой, причем матрица $\boldsymbol{B}$ определена однозначно с точностью до перестановки жордановых клеток на главной диагонали.*

§ 4.4. Диагонализируемые матрицы

Матрица $\boldsymbol{A} \in F^{n\times n}$ называется *диагонализируемой*, если она подобна диагональной матрице.

Предложение 5. *Квазидиагональная матрица диагонализируема тогда и только тогда, когда все ее диагональные блоки диагонализируемы.*

Доказательство. Пусть $\boldsymbol{A} = [\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m]$. Если существуют такие матрицы $\boldsymbol{S}_j$ $(j = 1, \ldots, m)$, что $\boldsymbol{S}_j^{-1}\boldsymbol{A}_j\boldsymbol{S}_j = \boldsymbol{D}_j$ — диагональные матрицы, то, полагая $\boldsymbol{S} = [\boldsymbol{S}_1, \ldots, \boldsymbol{S}_m]$, получим, что $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = [\boldsymbol{D}_1, \ldots, \boldsymbol{D}_m]$ – диагональная матрица.

Для доказательства обратного утверждения используем индукцию по m. Ясно, что достаточно рассмотреть случай $m = 2$. Пусть $\boldsymbol{A} = [\boldsymbol{A}_1, \boldsymbol{A}_2]$ и $\boldsymbol{A}_1 \in F^{k\times k}$, $\boldsymbol{A}_2 \in F^{n\times n}$. Пусть также $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D} = [\lambda_1, \ldots, \lambda_{k+n}]$ — диагональная матрица. Тогда $\boldsymbol{A}\boldsymbol{S} = \boldsymbol{S}\boldsymbol{D}$. Обозначим подматрицы j-го столбца матрицы $\boldsymbol{S}$ из первых k и последних n элементов через c_j и d_j соответственно. Столбцы $c_1, \ldots, c_{k+n}$ и $d_1, \ldots, d_{k+n}$ образуют в матрице $\boldsymbol{S}$ подматрицы $\boldsymbol{U}$ и $\boldsymbol{V}$ размеров $k\times(k+n)$ и $n\times(k+n)$ соответственно. Если $\mathsf{r}(\boldsymbol{U}) < k$ или $\mathsf{r}(\boldsymbol{V}) < n$, то $\mathsf{r}(\boldsymbol{S}) < k + n$, что противоречит обратимости $\boldsymbol{S}$. Следовательно, $\mathsf{r}(\boldsymbol{U}) = k$ и $\mathsf{r}(\boldsymbol{V}) = n$, так как эти матрицы имеют k и n строк соответственно. Таким образом, среди столбцов $c_1, \ldots, c_{k+n}$ имеется k, а среди $d_1, \ldots, d_{k+n}$ – n линейно независимых.

Поскольку для $j = 1, \ldots, k + n$ имеем $\boldsymbol{A}_1 c_j = \lambda_j c_j$, $\boldsymbol{A}_2 d_j = \lambda_j d_j$, заключаем, что матрицы $\boldsymbol{A}_1$ и $\boldsymbol{A}_2$ диагонализируемы. Предложение доказано.

Пусть $\boldsymbol{A}, \boldsymbol{B} \in F^{n\times n}$. Эти матрицы называются *перестановочными*, если $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{B}\boldsymbol{A}$.

Они называются *одновременно диагонализируемыми*, если существует обратимая матрица $\boldsymbol{S} \in F^{n\times n}$ такая, что матрицы $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S}$ и $\boldsymbol{S}^{-1}\boldsymbol{B}\boldsymbol{S}$ являются диагональными.

Теорема 31. *Пусть матрицы $\boldsymbol{A}, \boldsymbol{B} \in F^{n\times n}$ диагонализируемые. Они являются перестановочными тогда и только тогда, когда они одновременно диагонализируемые.*

Доказательство. Как хорошо известно, любые диагональные матрицы одного порядка перестановочны. Если матрицы $\boldsymbol{A}, \boldsymbol{B}$ одновременно диагонализируемы, то $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{D}_1$ и $\boldsymbol{S}^{-1}\boldsymbol{B}\boldsymbol{S} = \boldsymbol{D}_2$ – диагональные матрицы. Имеем $\boldsymbol{A}\boldsymbol{B} = \boldsymbol{S}\boldsymbol{D}_1\boldsymbol{S}^{-1}\boldsymbol{S}\boldsymbol{D}_2\boldsymbol{S}^{-1} = \boldsymbol{S}\boldsymbol{D}_1\boldsymbol{D}_2\boldsymbol{S}^{-1}$ и аналогично $\boldsymbol{B}\boldsymbol{A} = \boldsymbol{S}\boldsymbol{D}_2\boldsymbol{D}_1\boldsymbol{S}^{-1}$. Следовательно, $\boldsymbol{A}$ и $\boldsymbol{B}$ перестановочны.

Пусть теперь $\boldsymbol{AB} = \boldsymbol{BA}$. Диагонализируем матрицу $\boldsymbol{A}$ так, чтобы одинаковые собственные значения стояли рядом: $\boldsymbol{S}^{-1}\boldsymbol{AS} = [\lambda_1, \ldots, \lambda_1, \lambda_2, \ldots, \lambda_2, \ldots, \lambda_m, \ldots, \lambda_m] = \boldsymbol{A}'$. Положим $\boldsymbol{S}^{-1}\boldsymbol{BS} = \boldsymbol{B}' = (b_{ij})$. Из равенства $\boldsymbol{A}'\boldsymbol{B}' = \boldsymbol{B}'\boldsymbol{A}'$ выводим, считая, что λ_s и λ_t находятся соответственно в i-й и j-й строках матрицы $\boldsymbol{A}'$, что $\lambda_s b_{ij} = b_{ij}\lambda_t$. Поэтому при $s \neq t$ имеем $b_{ij} = 0$. Таким образом, матрица $\boldsymbol{B}'$ является квазидиагональной: $\boldsymbol{B}' = [\boldsymbol{B}_1, \ldots, \boldsymbol{B}_m]$. Такое же разбиение на блоки имеет матрица $\boldsymbol{A}' = [\boldsymbol{A}_1, \ldots, \boldsymbol{A}_m]$. Так как матрица $\boldsymbol{B}'$ диагонализируема, по предложению 5 каждая матрица $\boldsymbol{B}_j$ при $j = 1, \ldots, m$ диагонализируема: $\boldsymbol{T}_j^{-1}\boldsymbol{B}_j\boldsymbol{T}_j = \boldsymbol{D}_j$. Поскольку $\boldsymbol{A}_j = \lambda_j \boldsymbol{E}_{k_j}$, где k_j – порядок матрицы $\boldsymbol{A}_j$ (количество повторений λ_j на главной диагонали $\boldsymbol{A}'$), заключаем, что $\boldsymbol{T}_j^{-1}\boldsymbol{A}_j\boldsymbol{T}_j = \boldsymbol{A}_j$. Следовательно, квазидиагональная матрица $\boldsymbol{T} = [\boldsymbol{T}_1, \ldots, \boldsymbol{T}_m]$ диагонализует матрицы $\boldsymbol{A}'$ и $\boldsymbol{B}'$, и потому матрица $\boldsymbol{ST}$ диагонализует матрицы $\boldsymbol{A}$ и $\boldsymbol{B}$. Теорема доказана.

Семейство (конечное или бесконечное) матриц $\mathcal{S} \subseteq F^{n\times n}$ называется *коммутативным*, если для любых $\boldsymbol{A}, \boldsymbol{B} \in \mathcal{S}$ справедливо $\boldsymbol{AB} = \boldsymbol{BA}$.

Семейство матриц $\mathcal{S} \subseteq F^{n\times n}$ называется *одновременно диагонализируемым*, если существует такая невырожденная матрица $\boldsymbol{T} \in F^{n\times n}$, что для любой матрицы $\boldsymbol{A} \in \mathcal{S}$ матрица $\boldsymbol{T}^{-1}\boldsymbol{AT}$ является диагональной.

Теорема 32. *Любое семейство диагонализируемых матриц из $F^{n\times n}$ является коммутативным тогда и только тогда, когда оно одновременно диагонализируемо.*

Доказательство. Легко видеть, что если семейство матриц одновремнно диагонализируемо, то оно коммутативно (см. доказательство теоремы 31).

Пусть $\mathcal{S}$ – коммутативное семейство диагонализируемых матриц из F_n. Докажем индукцией по n, что оно одновремнно диагонализируемо. При $n = 1$ доказывать нечего. Пусть $n > 1$ — и для всех $1 \leq m < n$ утверждение доказано. Если $\mathcal{S}$ состоит только из скалярных матриц, доказывать также нечего. Будем считать, что $\mathcal{S}$ содержит матрицу $\boldsymbol{A}$, имеющую $2 \leq k \leq n$ различных собственных значений $\lambda_1, \ldots, \lambda_k$. Без ограничения общности предположим, что матрица $\boldsymbol{A}$ диагональная и на ее главной диагонали одинаковые собственные значения расположены рядом (см. доказательство теоремы 31). Поскольку $\boldsymbol{AB} = \boldsymbol{BA}$ для любой матрицы $\boldsymbol{B} \in \mathcal{S}$, как и в доказательстве теоремы 31, убеждаемся, что матрица $\boldsymbol{B}$ является квазидиагональной, причем разбиение на диагональные блоки определяется порядком и кратностями собственных значений матрицы $\boldsymbol{A}$ и поэтому является одним и тем же для всех матриц из $\mathcal{S}$. Порядок каждого блока не превосходит $n-1$. Так как все матрицы из $\mathcal{S}$ коммутируют между собой, блоки на соответствующих местах также коммутируют. Они диагонализируемы в силу предложения 5. Согласно предположению индукции блоки на соответствующих местах одновременно диагонализируемы с помощью матриц $\boldsymbol{T}_1, \ldots, \boldsymbol{T}_k$. Тогда прямая сумма этих матриц одновременно диагонализует все матрицы из $\mathcal{S}$.

Через F_n обозначим пространство n-элементных столбцов над полем F. Каждая матрица $\boldsymbol{A} \in F^{n\times n}$ определяет линейный оператор на пространстве F_n, сопоставляющий любому столбцу $x \in F_n$ произведение $\boldsymbol{A}x$.

Ненулевой вектор $x \in F_n$ называется *собственным* для матрицы $\boldsymbol{A}$, если $\boldsymbol{A}x = \lambda x$ для некоторого $\lambda \in F$; при этом λ называется *собственным значением* матрицы $\boldsymbol{A}$. Подпространство $W \subseteq F_n$ называется $\boldsymbol{A}$*-инвариантным* для матрицы $\boldsymbol{A} \in F^{n\times n}$, если $\boldsymbol{A}x \in W$ для любого $x \in W$. Подпространство $W \subseteq F_n$ называется $\mathcal{S}$*-инвариантным* для семейства матриц $\mathcal{S} \subseteq F^{n\times n}$, если $\boldsymbol{A}x \in W$ для любых $\boldsymbol{A} \in \mathcal{S}$ и $x \in W$.

Следующее утверждение хорошо известно из курса линейной алгебры.

Лемма 4. *Если характеристический многочлен матрицы $\boldsymbol{A}$ разлагается на линейные множители, то любое $\boldsymbol{A}$-инвариантное подпространство содержит собственный вектор этой матрицы.*

Предложение 6. *Пусть $\mathcal{S} \subseteq F^{n\times n}$ – коммутативное семейство матриц, у каждой из которых характеристический многочлен разлагается на линейные множители. Тогда в пространстве F_n существует общий собственный вектор для всех матриц из $\mathcal{S}$.*

Доказательство. Пусть W – $\mathcal{S}$-инвариантное подпространство из F_n наименьшей положительной размерности. Такое подпространство существует, так как F_n само является n-мерным $\mathcal{S}$-инвариантным подпространством. Покажем от противного, что любой ненулевой вектор из W является собственным вектором для любой матрицы из $\mathcal{S}$. Предположим, что не все ненулевые векторы из W являются собственным векторами для матрицы $\boldsymbol{A} \in \mathcal{S}$. По лемме существуют $\lambda \in F$ и ненулевой $x \in W$ такие, что $\boldsymbol{A}x = \lambda x$. Положим $W_0 = \{y \in W | \boldsymbol{A}y = \lambda y\}$. Тогда W_0 является $\boldsymbol{A}$-инвариантным и $0 < \mathsf{dim} W_0 < \mathsf{dim} W$. Покажем, что W_0 является $\boldsymbol{B}$-инвариантным для любой матрицы $\boldsymbol{B} \in \mathcal{S}$, отличной от $\boldsymbol{A}$. Так как $\boldsymbol{AB} = \boldsymbol{BA}$, для $y \in W_0$ имеем $\boldsymbol{A}(\boldsymbol{B}y) = (\boldsymbol{AB})y = (\boldsymbol{BA})y = = \boldsymbol{B}(\boldsymbol{A}y) = \boldsymbol{B}(\lambda y) = \lambda(\boldsymbol{B}y)$, откуда $\boldsymbol{B}y \in W_0$. Таким образом, W_0 является $\mathcal{S}$-инвариантным подпространством, что противоречит выбору W. Предложение доказано.

Предложение 7. *Пусть $\boldsymbol{A} \in F^{n\times n}$ – диагонализируемая матрица. Тогда для любого натурального $1 \le k \le n$ k-я ассоциированная с $\boldsymbol{A}$ матрица $\breve{\boldsymbol{A}}_k$ является диагонализируемой.*

Доказательство. Пусть $\boldsymbol{T}$ – невырожденная, $\boldsymbol{D}$ – диагональная матрицы такие, что $\boldsymbol{T}^{-1}\boldsymbol{AT} = \boldsymbol{D}$. Тогда по свойствам ассоциированных матриц имеем $\breve{\boldsymbol{T}}_k^{-1}\breve{\boldsymbol{A}}_k\breve{\boldsymbol{T}}_k = \breve{\boldsymbol{D}}_k$. Легко видеть, что $\breve{\boldsymbol{D}}_k$ – диагональная матрица, у которой на диагонали – всевозможные произведения диагональных элементов матрицы $\boldsymbol{D}$, взятых по k сомножителей. Предложение доказано.

§ 4.5. Матрицы, перестановочные с данной матрицей

Пусть поле F алгебраически замкнуто, $\boldsymbol{A} \in F^{n\times n}$. Какие матрицы удовлетворяют условию $\boldsymbol{AX} = \boldsymbol{XA}$, т. е. являются перестановочными с матрицей $\boldsymbol{A}$? Для ответа на этот вопрос приведем матрицу $\boldsymbol{A}$ в жордановой нормальной форме: для некоторой невырожденной матрицы $\boldsymbol{T} \in F^{n\times n}$ матрица $\boldsymbol{T}^{-1}\boldsymbol{AT} = \boldsymbol{J}$ является квазидиагональной матрицей с клетками Жордана на главной диагонали. Положим $\boldsymbol{X}_1 = \boldsymbol{T}^{-1}\boldsymbol{XT}$. Ясно, что $\boldsymbol{AX} = \boldsymbol{XA}$ тогда и только тогда, когда $\boldsymbol{JX}_1 = \boldsymbol{X}_1\boldsymbol{J}$. Будем решать последнее уравнение.

Сначала рассмотрим случай, когда матрица $\boldsymbol{J}$ является диагональной, т. е. все клетки Жордана одноэлементны. Будем считать, что одинаковые значения расположены рядом на главной диагонали. Если всего имеется m собственных значений с кратностями соответственно $k_1, \dots, k_m$ ($k_1 + \ldots + k_m = n$), то, как установлено в доказательстве теоремы 31, в этом случае матрица $\boldsymbol{X}_1$ является квазидиагональной матрицей с произвольными блоками порядков соответственно $k_1, \dots, k_m$ на главной диагонали.

Теперь предположим, что $\boldsymbol{J} = [\boldsymbol{J}_1, \dots, \boldsymbol{J}_s]$. Разобъем матрицу $\boldsymbol{X}_1$ на блоки в соответствии с порядками матриц $\boldsymbol{J}_1, \dots, \boldsymbol{J}_s$:

$$\boldsymbol{X}_1 = \begin{pmatrix} \boldsymbol{X}_{11} & \boldsymbol{X}_{12} & \dots & \boldsymbol{X}_{1s} \\ \boldsymbol{X}_{21} & \boldsymbol{X}_{22} & \dots & \boldsymbol{X}_{2s} \\ \dots & \dots & \dots & \dots \\ \boldsymbol{X}_{s1} & \boldsymbol{X}_{s2} & \dots & \boldsymbol{X}_{ss} \end{pmatrix}.$$

Равенство $\boldsymbol{J}\boldsymbol{X}_1 = \boldsymbol{X}_1\boldsymbol{J}$ равносильно совокупности равенств $\boldsymbol{J}_p\boldsymbol{X}_{pq} = \boldsymbol{X}_{pq}\boldsymbol{J}_q$ для всех $p, q = 1, \dots, s$. Предположим, что клетка Жордана $\boldsymbol{J}_p$ имеет порядок k и относится к собственному значению ρ, а клетка Жордана $\boldsymbol{J}_q$ имеет порядок m и относится к собственному значению σ. Тогда матрица $\boldsymbol{X}_{pq} = (\xi_{ij})$ имеет размеры $k \times m$. Запишем равенство $\boldsymbol{J}_p\boldsymbol{X}_{pq} = \boldsymbol{X}_{pq}\boldsymbol{J}_q$:

$$\begin{pmatrix} \rho & 1 & 0 & \dots & 0 \\ 0 & \rho & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \rho \end{pmatrix} \begin{pmatrix} \xi_{11} & \xi_{12} & \dots & \xi_{1m} \\ \xi_{21} & \xi_{22} & \dots & \xi_{2m} \\ \dots & \dots & \dots & \dots \\ \xi_{k1} & \xi_{k2} & \dots & \xi_{km} \end{pmatrix} =$$

$$= \begin{pmatrix} \xi_{11} & \xi_{12} & \dots & \xi_{1m} \\ \xi_{21} & \xi_{22} & \dots & \xi_{2m} \\ \dots & \dots & \dots & \dots \\ \xi_{k1} & \xi_{k2} & \dots & \xi_{km} \end{pmatrix} \begin{pmatrix} \sigma & 1 & 0 & \dots & 0 \\ 0 & \sigma & 1 & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & 1 \\ 0 & 0 & 0 & \dots & \sigma \end{pmatrix}.$$

Приравниваем элементы с индексами ij в левой и правой частях этого равенства, рассматривая возможные случаи для индексов.

1. $i < k$, $j > 1$. $\rho\xi_{ij} + \xi_{i+1,j} = \xi_{i,j-1} + \sigma\xi_{ij}$.
2. $i < k$, $j = 1$. $\rho\xi_{i1} + \xi_{i+1,1} = \sigma\xi_{i1}$.
3. $i = k$, $j > 1$. $\rho\xi_{kj} = \xi_{k,j-1} + \sigma\xi_{kj}$.
4. $i = k$, $j = 1$. $\rho\xi_{k1} = \sigma\xi_{k1}$.

Рассмотрим случай, когда $\rho \neq \sigma$. Из равенства 4 следует $\xi_{k1} = 0$. Из равенств 3 получаем $\xi_{kj} = 0$ при $j = 2, \dots, m$. Из 2 получаем

$\xi_{i1} = 0$ при $i = k-1, \dots, 1$. Наконец, из равенств 1 следует $\xi_{ij} = 0$ при $2 \le i < k$, $1 < j \le m$. Таким образом, если $\rho \ne \sigma$, то $\boldsymbol{X}_{pq} = \boldsymbol{O}$.

Рассмотрим случай, когда $\rho = \sigma$. Тогда равенство 4 выполняется всегда, а равенства 1, 2 и 3 превращаются соответственно в:

1'. $i < k$, $j > 1$. $\xi_{i+1,j} = \xi_{i,j-1}$.

2'. $i < k$, $j = 1$. $\xi_{i+1,1} = 0$.

3'. $i = k$, $j > 1$. $\xi_{k,j-1} = 0$.

Если $k = m$, то матрица $\boldsymbol{X}_{pq}$ – квадратная, верхнетреугольная и все элементы на главной диагонали равны между собой, все элементы каждой диагонали выше главной также равны между собой. Матрица $\boldsymbol{X}_{pq}$ определяется k независимыми параметрами.

Если $k < m$, то матрица описанного вида находится в правых k столбцах матрицы $\boldsymbol{X}_{pq}$, а остальные ее столбцы нулевые.

Если $k > m$, то матрица описанного вида находится в первых k строках матрицы $\boldsymbol{X}_{pq}$, а остальные ее строки нулевые.

Рассмотрим пример. Пусть

$$\boldsymbol{J} = [\boldsymbol{J}_1, \boldsymbol{J}_2], \text{ и } \boldsymbol{J}_1 = \begin{pmatrix} \rho & 1 & 0 \\ 0 & \rho & 1 \\ 0 & 0 & \rho \end{pmatrix}, \boldsymbol{J}_2 = \begin{pmatrix} \rho & 1 \\ 0 & \rho \end{pmatrix}.$$

Тогда

$$\boldsymbol{X}_1 = \begin{pmatrix} \alpha & \beta & \gamma & \delta & \epsilon \\ 0 & \alpha & \beta & 0 & \delta \\ 0 & 0 & \alpha & 0 & 0 \\ 0 & \zeta & \eta & \theta & \kappa \\ 0 & 0 & \zeta & 0 & \theta \end{pmatrix},$$

где $\alpha, \beta, \gamma, \delta, \epsilon, \zeta, \eta, \theta, \kappa \in F$ – независимые параметры.

Задачи для самостоятельного решения

1. Найти естественную и вторую естественную нормальные формы для следующих матриц над полем действительных чисел:

$$\text{а) } \begin{pmatrix} 10 & 29 & 53 & -38 \\ 0 & 0 & 1 & -2 \\ -10 & -28 & -53 & 38 \\ -10 & -28 & -53 & 37 \end{pmatrix}; \text{ б) } \begin{pmatrix} -1 & 1 & 1 & 1 \\ 0 & -1 & -1 & -2 \\ 0 & 0 & 5 & 10 \\ 0 & 0 & -5 & -9 \end{pmatrix}.$$

2. Найти естественную и вторую естественную нормальные формы для следующих матриц над полем рациональных чисел:

а) $\begin{pmatrix} -2 & 0 & 2 & 1 \\ 0 & 1 & 1 & 0 \\ -2 & 0 & 2 & 3 \\ 2 & 0 & -2 & -2 \end{pmatrix}$; б) $\begin{pmatrix} -3 & -9 & -15 & -28 \\ 0 & -1 & -1 & -2 \\ -3 & -9 & -15 & -26 \\ 3 & 9 & 15 & 27 \end{pmatrix}$.

3. Найти жорданову нормальную форму для следующих матриц:

а) $\begin{pmatrix} -1 & -4 & -8 & -11 & -4 & -8 \\ 0 & 1 & -1 & -1 & -1 & -1 \\ 0 & 0 & -1 & -1 & -1 & -1 \\ 1 & 1 & 3 & 6 & 1 & 2 \\ 1 & -1 & 0 & 1 & 1 & -1 \\ -1 & -1 & 0 & -1 & 2 & 4 \end{pmatrix}$;

б) $\begin{pmatrix} 4 & -1 & 5 & 11 & 12 & 11 \\ 4 & 5 & 11 & 15 & 7 & 11 \\ -3 & -3 & -8 & -9 & -3 & -6 \\ 0 & 0 & 0 & -2 & -7 & -7 \\ 1 & -3 & -2 & -1 & 1 & -3 \\ -1 & 3 & 2 & 1 & 4 & 8 \end{pmatrix}$.

4. Убедиться, что матрицы $\boldsymbol{A}$ и $\boldsymbol{B}$ подобны, и найти трансформирующую матрицу:

а) $\boldsymbol{A} = \begin{pmatrix} 17 & -6 \\ 45 & -16 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 14 & -60 \\ 3 & -13 \end{pmatrix}$;

б) $\boldsymbol{A} = \begin{pmatrix} 8 & 7 \\ -9 & 15 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} -24 & -57 \\ 23 & 47 \end{pmatrix}$.

5. Выяснить, являются ли подобными следующие матрицы, и в случае подобия найти трансформирующую матрицу:

$\boldsymbol{A} = \begin{pmatrix} 4 & 6 & -15 \\ 1 & 3 & -5 \\ 1 & 2 & -4 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 1 & -3 & 3 \\ -2 & -6 & 13 \\ -1 & -4 & 8 \end{pmatrix}$,

$\boldsymbol{C} = \begin{pmatrix} -13 & -70 & 119 \\ -4 & -19 & 34 \\ -4 & -20 & 35 \end{pmatrix}$.

6. Найти все матрицы, перестановочные с данной матрицей $\boldsymbol{A}$:

а) $\boldsymbol{A} = \begin{pmatrix} 1 & 4 & 4 & 4 & 3 & 5 \\ 1 & 3 & 4 & 5 & 3 & 4 \\ 0 & -1 & 0 & -1 & -2 & -2 \\ 0 & -1 & -1 & 0 & 0 & -1 \\ 1 & 2 & 4 & 5 & 4 & 5 \\ -1 & -1 & -3 & -4 & -2 & -2 \end{pmatrix}$;

б) $$\boldsymbol{A} = \begin{pmatrix} 0 & 2 & 0 & -3 & 0 & 0 \\ -1 & 1 & -2 & -4 & -2 & -3 \\ 1 & 0 & 3 & 3 & 0 & 1 \\ 0 & -1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 1 & 1 & 1 & 1 & 3 \end{pmatrix}.$$

Глава 5

Характеристический и минимальный многочлены матрицы

§ 5.1. Характеристический многочлен

Напомним, что *характеристическим многочленом* матрицы $\boldsymbol{A} \in F^{n\times n}$ называется определитель ее характеристической матрицы $x\boldsymbol{E}_n - \boldsymbol{A}$, который обозначается через $\chi_{\boldsymbol{A}}(x)$.

Как показывает следующее утверждение, коэффициенты характеристического многочлена матрицы выражаются через ее главные миноры.

Предложение 8. *Пусть* $\boldsymbol{A} \in F^{n\times n}$. *Тогда*

$$\chi_{\boldsymbol{A}}(x) = x^n - p_1 x^{n-1} - p_2 x^{n-2} - \ldots - p_{n-1}x - p_n, \tag{1}$$

где $p_k = (-1)^{k-1} \sum_{1\le j_1<\ldots<j_k\le n} \boldsymbol{A}\begin{pmatrix} j_1 & j_2 & \ldots & j_k \\ j_1 & j_2 & \ldots & j_k \end{pmatrix}$.

Доказательство. Пусть $\boldsymbol{A} = (a_{ij})$. Тогда

$$\chi_{\boldsymbol{A}}(x) = \begin{vmatrix} x - a_{11} & -a_{12} & \dots & -a_{1n} \\ -a_{21} & x - a_{22} & \dots & -a_{2n} \\ \dots & \dots & \dots & \dots \\ -a_{n1} & -a_{n2} & \dots & x - a_{nn} \end{vmatrix}.$$

Слагаемые, содержащие x^{n-k} при $k \geq 1$, получаются при раскрытии скобок в слагаемых определителя, содержащих произведение $(x - a_{i_1 i_1}) \dots (x - a_{i_{n-k} i_{n-k}})$ при всевозможных $1 \leq i_1 < \dots < < i_{n-k} \leq n$. Зафиксируем значения $1 \leq i_1 < \dots < i_{n-k} \leq n$ и через $1 \leq j_1 < \dots < j_k \leq n$ обозначим индексы из дополнения множества $\{i_1, \dots, i_{n-k}\}$ до $\{1, \dots, n\}$. Коэффициент при x^{n-k} в сумме всех слагаемых, содержащих произведение $(x - a_{i_1 i_1}) \dots (x - a_{i_{n-k} i_{n-k}})$, равен минору $(-1)^k \boldsymbol{A} \begin{pmatrix} j_1 & j_2 & \dots & j_k \\ j_1 & j_2 & \dots & j_k \end{pmatrix}$. Предложение доказано.

В частности, $p_1 = \mathsf{tr}(\boldsymbol{A})$, $p_n = (-1)^{n-1} |\boldsymbol{A}|$.

Следующее утверждение характеризует корни характеристического многочлена матрицы, являющейся значением некоторого многочлена от другой матрицы.

Теорема 33. *Пусть F — алгебраически замкнутое поле, $\boldsymbol{A} \in F^{n \times n}$, $g(x) \in F[x]$, $\lambda_1, \dots, \lambda_n$ — все корни характеристического многочлена $\chi_{\boldsymbol{A}}(x)$ матрицы $\boldsymbol{A}$ с учетом кратности. Тогда $g(\lambda_1), \dots, g(\lambda_n)$ — все корни характеристического многочлена $\chi_{g(\boldsymbol{A})}(x)$ матрицы $g(\boldsymbol{A})$ с учетом кратности.*

Доказательство. Пусть $\chi_{\boldsymbol{A}}(x) = \prod_{j=1}^{n} (x - \lambda_j)$, $g(x) = a_0 \prod_{\ell=1}^{m} (x - \mu_\ell)$. Тогда $g(\boldsymbol{A}) = a_0 \prod_{\ell=1}^{m} (\boldsymbol{A} - \mu_\ell \boldsymbol{E}_n)$. Перейдем к определителям:

$$|g(\boldsymbol{A})| = |a_0 \textstyle\prod_{\ell=1}^{m} (\boldsymbol{A} - \mu_\ell \boldsymbol{E}_n)| = a_0^n (-1)^m \displaystyle\prod_{\ell=1}^{m} |\mu_\ell \boldsymbol{E}_n - \boldsymbol{A}| =$$

$$= (-1)^m a_0^n \prod_{\ell=1}^{m} \chi_{g(\boldsymbol{A})}(\mu_\ell) = (-1)^m a_0^n \prod_{\ell=1}^{m} \prod_{j=1}^{n} (\mu_\ell - \lambda_j) =$$

$$= \prod_{j=1}^{n} a_0 \prod_{\ell=1}^{m} (\lambda_j - \mu_\ell) = \prod_{j=1}^{n} g(\lambda_j),$$

откуда $|g(\boldsymbol{A})| = \prod_{j=1}^{n} g(\lambda_j)$. В эту формулу подставим вместо многочлена $g(x)$ многочлен $y-g(x)$. Получим $|y\boldsymbol{E}_n - g(\boldsymbol{A})| = \prod_{j=1}^{n}(y- -g(\lambda_j))$. Таким образом, $\chi_{g(\boldsymbol{A})}(x) = \prod_{j=1}^{n}(x - g(\lambda_j))$, что и требуется доказать.

Зафиксируем матрицу $\boldsymbol{A} \in F^{n\times n}$. Обозначим через $\boldsymbol{B}(x)$ присоединенную к характеристической матрице $x\boldsymbol{E}_n - \boldsymbol{A}$. Так как

$$(x\boldsymbol{E}_n - \boldsymbol{A})\boldsymbol{B}(x) = \chi_{\boldsymbol{A}}(x)\boldsymbol{E}_n, \tag{2}$$

заключаем, что многочлен $\chi_{\boldsymbol{A}}(x)\boldsymbol{E}_n$ делится на $x\boldsymbol{E}_n - \boldsymbol{A}$. Отсюда в силу следствия 5 вытекает, что $\chi_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{O}$. Получаем еще одно доказательство теоремы Гамильтона–Кэли.

Выразим коэффициенты $\boldsymbol{B}(x)$ через характеристический многочлен $\chi_{\boldsymbol{A}}(x)$. Из равенства (2) следует $(x\boldsymbol{E}_n - \boldsymbol{A})(x^{n-1}\boldsymbol{E}_n + \boldsymbol{B}_1 x^{n-2} + \ldots + \boldsymbol{B}_{n-1}) = (x^n - p_1 x^{n-1} - p_2 x^{n-2} - \ldots - p_{n-1}x - p_n)\boldsymbol{E}_n$. Приравнивая коэффициенты при соответствующих степенях x, получаем $\boldsymbol{B}_1 - \boldsymbol{A} = -p_1\boldsymbol{E}_n$, $\boldsymbol{B}_2 - \boldsymbol{A}\boldsymbol{B}_1 = -p_2\boldsymbol{E}_n, \ldots, \boldsymbol{B}_{n-1} - \boldsymbol{A}\boldsymbol{B}_{n-2} = -p_{n-1}\boldsymbol{E}_n$, $-\boldsymbol{A}\boldsymbol{B}_{n-1} = -p_n\boldsymbol{E}_n$. Таким образом, для вычисления матриц $\boldsymbol{B}_i$ имеем рекуррентные соотношения

$$\boldsymbol{B}_0 = \boldsymbol{E}_n,\ \boldsymbol{B}_k = \boldsymbol{A}\boldsymbol{B}_{k-1} - p_k\boldsymbol{E}_n \ (k = 1, \ldots, n-1), \tag{3}$$

при этом $\boldsymbol{A}\boldsymbol{B}_{n-1} = p_n\boldsymbol{E}_n$. Если $|\boldsymbol{A}| \neq 0$, то $\boldsymbol{A}^{-1} = \frac{1}{p_n}\boldsymbol{B}_{n-1}$, так как $p_n = (-1)^{n-1}|\boldsymbol{A}| \neq 0$.

Если λ – собственное значение матрицы $\boldsymbol{A}$, то

$$(\lambda\boldsymbol{E}_n - \boldsymbol{A})\boldsymbol{B}(\lambda) = \chi_{\boldsymbol{A}}(\lambda)\boldsymbol{E}_n = \boldsymbol{O},$$

т. е. $\boldsymbol{B}(\lambda)$ состоит из столбцов – собственных векторов матрицы $\boldsymbol{A}$, относящихся к собственному значению λ.

Пусть $\boldsymbol{A}, \boldsymbol{B} \in F^{n\times n}$. Если по крайней мере одна из этих матриц невырожденная, то матрицы $\boldsymbol{AB}$ и $\boldsymbol{BA}$ подобны и потому их характеристические многочлены равны. Например, если $\boldsymbol{A}$ невырожденная, то $\boldsymbol{BA} = \boldsymbol{A}^{-1}(\boldsymbol{AB})\boldsymbol{A}$. Если обе матрицы вырожденные, то $\boldsymbol{AB}$ и $\boldsymbol{BA}$ не обязательно подобны. В качестве примера можно взять матрицы $\boldsymbol{A}, \boldsymbol{B} \in F^{n\times n}$, для которых $\boldsymbol{AB} = \boldsymbol{O}$ и $\boldsymbol{BA} \neq \boldsymbol{O}$.

В общем случае справедливо следующее утверждение.

Теорема 34. *Пусть* $\boldsymbol{A} \in F^{m\times n}$, $\boldsymbol{B} \in F^{n\times m}$ *и* $m \leq n$. *Тогда* $\chi_{\boldsymbol{BA}}(x) = x^{n-m}\chi_{\boldsymbol{AB}}(x)$.

Доказательство. Для блочных матриц порядка $m+n$ имеем $\begin{pmatrix} \boldsymbol{AB} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix}\begin{pmatrix} \boldsymbol{E}_m & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{E}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{AB} & \boldsymbol{ABA} \\ \boldsymbol{B} & \boldsymbol{BA} \end{pmatrix}$ и $\begin{pmatrix} \boldsymbol{E}_m & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{E}_n \end{pmatrix}\begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{BA} \end{pmatrix} = \begin{pmatrix} \boldsymbol{AB} & \boldsymbol{ABA} \\ \boldsymbol{B} & \boldsymbol{BA} \end{pmatrix}$. Таким образом, $\begin{pmatrix} \boldsymbol{AB} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix}\begin{pmatrix} \boldsymbol{E}_m & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{E}_n \end{pmatrix} = \begin{pmatrix} \boldsymbol{E}_m & \boldsymbol{A} \\ \boldsymbol{O} & \boldsymbol{E}_n \end{pmatrix}\begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{BA} \end{pmatrix}$, поэтому матрицы $\begin{pmatrix} \boldsymbol{AB} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{O} \end{pmatrix}$ и $\begin{pmatrix} \boldsymbol{O} & \boldsymbol{O} \\ \boldsymbol{B} & \boldsymbol{BA} \end{pmatrix}$ подобны. Следовательно, их характеристические многочлены равны, т. е. $\begin{vmatrix} x\boldsymbol{E}_m - \boldsymbol{AB} & \boldsymbol{O} \\ -\boldsymbol{B} & x\boldsymbol{E}_n \end{vmatrix} = \begin{vmatrix} x\boldsymbol{E}_m & \boldsymbol{O} \\ -\boldsymbol{B} & x\boldsymbol{E}_n - \boldsymbol{BA} \end{vmatrix}$. Так как $\begin{vmatrix} x\boldsymbol{E}_m - \boldsymbol{AB} & \boldsymbol{O} \\ -\boldsymbol{B} & x\boldsymbol{E}_n \end{vmatrix} = x^n|x\boldsymbol{E}_m - \boldsymbol{AB}| = x^n\chi_{\boldsymbol{AB}}(x)$ и $\begin{vmatrix} x\boldsymbol{E}_m & \boldsymbol{O} \\ -\boldsymbol{B} & x\boldsymbol{E}_n - \boldsymbol{BA} \end{vmatrix} = x^m|x\boldsymbol{E}_n - \boldsymbol{BA}| = x^m\chi_{\boldsymbol{BA}}(x)$, с учетом неравенства $m \geq n$ получаем требуемое.

Из теоремы 34 непосредственно получается

Следствие 8. *Если* $\boldsymbol{A}, \boldsymbol{B} \in F^{n\times n}$, *то* $\chi_{\boldsymbol{BA}}(x) = \chi_{\boldsymbol{AB}}(x)$.

§ 5.2. Метод Д. К. Фаддеева

Этот метод служит для одновременного определения скалярных коэффициентов $p_1, \ldots, p_n$ характеристического многочлена и матричных коэффициентов $\boldsymbol{B}_1, \ldots, \boldsymbol{B}_{n-1}$ присоединенной x-матрицы $\boldsymbol{B}(x)$ для матрицы $\boldsymbol{A} \in F^{n\times n}$ над алгебраически замкнутым полем F характеристики 0. Приведем описание алгоритма.

$\boldsymbol{A}_1 = \boldsymbol{A}, p_1 = \mathsf{tr}(\boldsymbol{A}_1), \boldsymbol{B}_1 = \boldsymbol{A}_1 - p_1\boldsymbol{E}_n$

$\boldsymbol{A}_2 = \boldsymbol{AB}_1, p_2 = \frac{1}{2}\mathsf{tr}(\boldsymbol{A}_2), \boldsymbol{B}_2 = \boldsymbol{A}_2 - p_2\boldsymbol{E}_n$

....

$\boldsymbol{A}_{n-1} = \boldsymbol{AB}_{n-2}, p_{n-1} = \frac{1}{n-1}\mathsf{tr}(\boldsymbol{A}_{n-1}), \boldsymbol{B}_{n-1} = \boldsymbol{A}_{n-1} - p_{n-1}\boldsymbol{E}_n$

$\boldsymbol{A}_n = \boldsymbol{AB}_{n-1}, p_n = \frac{1}{n}\mathsf{tr}(\boldsymbol{A}_n), \boldsymbol{B}_n = \boldsymbol{A}_n - p_n\boldsymbol{E}_n = \boldsymbol{O}$.

Обоснование алгоритма использует формулу Ньютона для симметрических многочленов от переменных $x_1, \ldots, x_n$. Обозначим через s_m симметрический многочлен $x_1^m + \ldots + x_n^m$ и через σ_k $(1 \leq \leq k \leq n)$ элементарные симметрические многочлены от переменных $x_1, \ldots, x_n$. Тогда при $1 < k \leq n$ справедлива формула Ньютона

$$s_k - s_{k-1}\sigma_1 + s_{k-2}\sigma_2 + \ldots + (-1)^{k-1}s_1\sigma_{k-1} + (-1)^k k\sigma_k = 0. \qquad (1)$$

Пусть $\lambda_1, \ldots, \lambda_n$ – все корни характеристического многочлена $\chi_{\boldsymbol{A}}(x)$ с учетом кратности. По предложению 8, $p_1 = \mathsf{tr}(\boldsymbol{A})$. Согласно теореме 33 для натурального числа m имеем

$$\chi_{\boldsymbol{A}^m}(x) = (x - \lambda_1^m) \ldots (x - \lambda_n^m),$$

поэтому $\mathsf{tr}(\boldsymbol{A}^m) = \lambda_1^m + \ldots + \lambda_n^m = s_m(\lambda_1, \ldots, \lambda_n)$. Для краткости обозначим $s_m(\lambda_1, \ldots, \lambda_n)$ через s_m. Согласно обобщенной теореме Виета, примененной к многочлену $\chi_{\boldsymbol{A}}(x)$, заключаем, что

$$p_m = (-1)^{m-1}\sigma_m(\lambda_1, \ldots, \lambda_n).$$

Двигаясь по формулам алгоритма сверху вниз, получаем
$\boldsymbol{A}_2 = \boldsymbol{A}(\boldsymbol{A} - p_1\boldsymbol{E}_n) = \boldsymbol{A}^2 - p_1\boldsymbol{A}$,
$\boldsymbol{A}_3 = \boldsymbol{A}^3 - p_1\boldsymbol{A}^2 - p_2\boldsymbol{A}$, ...,
$\boldsymbol{A}_m = \boldsymbol{A}^m - p_1\boldsymbol{A}^{m-1} - \ldots - p_{m-1}\boldsymbol{A}$ $(4 \leq m \leq n)$.

Таким образом, $\mathsf{tr}(\boldsymbol{A}_m) = \mathsf{tr}(\boldsymbol{A}^m) - p_1\mathsf{tr}(\boldsymbol{A}^{m-1}) - \ldots - p_{m-1}\mathsf{tr}(\boldsymbol{A}) = s_m - \sigma_1 s_{m-1} + \sigma_2 s_{m-2} + \ldots + (-1)^{m-1}\sigma_{m-1}s_1$. Согласно равенству (1) имеем

$$s_m - \sigma_1 s_{m-1} + \sigma_2 s_{m-2} + \ldots + (-1)^{m-1}\sigma_{m-1}s_1 = (1)^{m-1}m\sigma_m = mp_m.$$

Следовательно, $p_m = \frac{1}{m}\mathsf{tr}(\boldsymbol{A}_m)$.

Формулы для матричных коэффициентов $\boldsymbol{B}_m$ из алгоритма соответствуют формулам (3) из § 5.1.

Заметим, что если матрица $\boldsymbol{A} \in F^{n\times n}$ для некоторого поля F характеристики 0, то метод Д. К. Фаддеева можно применять, так как любое поле можно включить как подполе в алгебраически замкнутое поле, а все вычисления метода производятся в поле F.

Рассмотрим пример.

$\boldsymbol{A} = \begin{pmatrix} 2 & -1 & 1 & 2 \\ 0 & 1 & 1 & 0 \\ -1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$, $p_1 = \mathsf{tr}(\boldsymbol{A}) = 4$, $\boldsymbol{B}_1 = \boldsymbol{A} - 4\boldsymbol{E}_4 =$

$$= \begin{pmatrix} -2 & -1 & 1 & 2 \\ 0 & -3 & 1 & 0 \\ -1 & 1 & -3 & 1 \\ 1 & 1 & 1 & -4 \end{pmatrix}; \ \boldsymbol{A}_2 = \boldsymbol{A}\boldsymbol{B}_1 = \begin{pmatrix} -3 & 4 & 0 & -3 \\ -1 & -2 & -2 & 1 \\ 2 & 0 & -2 & -5 \\ -3 & -3 & -1 & 3 \end{pmatrix},$$

$$p_2 = \tfrac{1}{2}\mathsf{tr}(\boldsymbol{A}_2) = -2, \ \boldsymbol{B}_2 = \boldsymbol{A}_2 + 2\boldsymbol{E}_4 = \begin{pmatrix} -1 & 4 & 0 & -3 \\ -1 & 0 & -2 & 1 \\ 2 & 0 & 0 & -5 \\ -3 & -3 & -1 & 5 \end{pmatrix};$$

$$\boldsymbol{A}_3 = \boldsymbol{A}\boldsymbol{B}_2 = \begin{pmatrix} -5 & 2 & 0 & 2 \\ 1 & 0 & -2 & -4 \\ -1 & -7 & -3 & 4 \\ 0 & 4 & -2 & -7 \end{pmatrix}, \ p_3 = \tfrac{1}{3}\mathsf{tr}(\boldsymbol{A}_3) = -5, \ \boldsymbol{B}_3 =$$

$$= \boldsymbol{A}_3 + 5\boldsymbol{E}_4 = \begin{pmatrix} 0 & 2 & 0 & 2 \\ 1 & 5 & -2 & -4 \\ -1 & -7 & 2 & 4 \\ 0 & 4 & -2 & -2 \end{pmatrix}; \ \boldsymbol{A}_4 = \boldsymbol{A}\boldsymbol{B}_3 =$$

$$= \begin{pmatrix} -2 & 0 & 0 & 0 \\ 0 & -2 & 0 & 0 \\ 0 & 0 & -2 & 0 \\ 0 & 0 & 0 & -2 \end{pmatrix}, \ p_4 = \tfrac{1}{4}\mathsf{tr}(\boldsymbol{A}_3) = -2, \ \boldsymbol{B}_4 = \boldsymbol{A}_4 + 2\boldsymbol{E}_4 = \boldsymbol{O};$$

$$\chi_{\boldsymbol{A}}(x) = x^4 - 4x^3 + 2x^2 + 5x + 2, \ |\boldsymbol{A}| = 2, \ \boldsymbol{A}^{-1} = \tfrac{1}{p_4}\boldsymbol{B}_3.$$

§ 5.3. Минимальный многочлен

Минимальным многочленом матрицы $\boldsymbol{A} \in F^{n\times n}$ называется многочлен наименьшей степени со старшим коэффициентом 1, аннулирующий матрицу $\boldsymbol{A}$. Обозначим его через $\mu_{\boldsymbol{A}}(x)$.

Для матрицы $\boldsymbol{A} \in F^{n\times n}$ обозначим присоединенную матрицу к $x\boldsymbol{E}_n - \boldsymbol{A}$ через $\boldsymbol{B}(x)$. Тогда $\boldsymbol{B}(x) = D_{n-1}\boldsymbol{C}(x)$, где D_{n-1} – наибольший общий делитель миноров порядка $n-1$ матрицы $x\boldsymbol{E}_n - \boldsymbol{A}$, а $\boldsymbol{C}(x)$ – матрица, называемая *приведенной присоединенной*. Легко понять, что все элементы $\boldsymbol{C}(x)$ взаимно просты в совокупности.

Предложение 9. *Минимальный многочлен $\mu_{\boldsymbol{A}}(x)$ матрицы $\boldsymbol{A} \in F^{n\times n}$ совпадает с многочленом $i_n(x)$ – n-м инвариантным множителем характеристической матрицы $x\boldsymbol{E}_n - \boldsymbol{A}$.*

Доказательство. Убедимся, что $i_n(\boldsymbol{A}) = \boldsymbol{O}$. Согласно (2) имеем $\chi_{\boldsymbol{A}}(x)\boldsymbol{E}_n = (x\boldsymbol{E}_n - \boldsymbol{A})\boldsymbol{B}(x)$. Так как $\chi_{\boldsymbol{A}}(x) = D_n$ и $D_n = D_{n-1}i_n(x)$, получаем $D_{n-1}i_n(x)\boldsymbol{E}_n = (x\boldsymbol{E}_n - \boldsymbol{A})D_{n-1}\boldsymbol{C}(x)$. Сократив на D_{n-1}, имеем $i_n(x)\boldsymbol{E}_n = (x\boldsymbol{E}_n - \boldsymbol{A})\boldsymbol{C}(x)$, откуда следует, что $i_n(\boldsymbol{A}) = \boldsymbol{O}$.

Таким образом, $\mu_{\boldsymbol{A}}(x)$ делит $i_n(x)$. Пусть $i_n(x) = f(x)\mu_{\boldsymbol{A}}(x)$ для некоторого многочлена $f(x)$. Поскольку $\mu_{\boldsymbol{A}}(\boldsymbol{A}) = \boldsymbol{O}$, справедливо равенство $\mu_{\boldsymbol{A}}(x)\boldsymbol{E}_n = (x\boldsymbol{E}_n - \boldsymbol{A})\boldsymbol{S}(x)$ для некоторой x-матрицы $\boldsymbol{S}(x)$. Отсюда следует $(x\boldsymbol{E}_n - \boldsymbol{A})\boldsymbol{C}(x) = f(x)(x\boldsymbol{E}_n - \boldsymbol{A})\boldsymbol{S}(x)$. Сокращая на регулярный многочлен $x\boldsymbol{E}_n - \boldsymbol{A}$, получаем $\boldsymbol{C}(x) = f(x)\boldsymbol{S}(x)$. Так как все элементы $\boldsymbol{C}(x)$ взаимно просты в совокупности, заключаем, что $f(x) = 1$. Предложение доказано.

Следствие 9. *Для любой матрицы $\boldsymbol{A} \in F^{n\times n}$ некоторая степень ее минимального многочлена $\mu_{\boldsymbol{A}}(x)$ делится на характеристический многочлен $\chi_{\boldsymbol{A}}(x)$. В частности, $\mu_{\boldsymbol{A}}(x)$ и $\chi_{\boldsymbol{A}}(x)$ имеют одни и те же корни.*

Доказательство. Из равенства

$$\mu_{\boldsymbol{A}}(x)\boldsymbol{E}_n = (x\boldsymbol{E}_n - \boldsymbol{A})\boldsymbol{C}(x), \tag{1}$$

переходя к определителям, получаем $\mu_{\boldsymbol{A}}(x)^n = \chi_{\boldsymbol{A}}(x)|\boldsymbol{C}(x)|$, откуда следует требуемое.

Для любого собственного значения λ матрицы $\boldsymbol{A}$ справедливо $\boldsymbol{C}(\lambda) \neq \boldsymbol{O}$, так как в противном случае все элементы $\boldsymbol{C}(x)$ делятся на $x - \lambda$. Из равенства $(\lambda\boldsymbol{E}_n - \boldsymbol{A})\boldsymbol{C}(\lambda) = \boldsymbol{O}$ следует, что ненулевые столбцы матрицы $\boldsymbol{C}(\lambda)$ являются собственными векторами матрицы $\boldsymbol{A}$.

Запишем $\mu_{\boldsymbol{A}}(x) = x^k + m_1x^{k-1} + \ldots + m_{k-1}x + m_k$ и положим $\Psi(x, y) = \dfrac{\mu_{\boldsymbol{A}}(x) - \mu_{\boldsymbol{A}}(y)}{x - y}$. Имеем

$$\begin{aligned}\Psi(x, y) = \frac{x^k - y^k + m_1(x^{k-1} - y^{k-1}) + \ldots + m_{k-1}(x - y)}{x - y} = x^{k-1} + \\ + x^{k-2}y + \ldots + xy^{k-2} + y^{k-1} + m_1(x^{k-2} + x^{k-3}y + \ldots + xy^{k-3} + y^{k-2}) + \\ + \ldots + m_{k-1} = x^{k-1} + x^{k-2}(y + m_1) + x^{k-3}(y^2 + m_1y + m_2) + \ldots + \\ + y^{k-1} + m_1y^{k-2} + \ldots + m_{k-1}.\end{aligned}$$

Далее, подставив в равенство $\mu_{\boldsymbol{A}}(x) - \mu_{\boldsymbol{A}}(y) = (x-y)\Psi(x,y)$ матрицу $x\boldsymbol{E}_n$ вместо x и $\boldsymbol{A}$ вместо y, получим

$$\mu_{\boldsymbol{A}}(x)\boldsymbol{E}_n - \mu_{\boldsymbol{A}}(\boldsymbol{A}) = (x\boldsymbol{E}_n - \boldsymbol{A})\Psi(x\boldsymbol{E}_n, \boldsymbol{A}),$$

откуда следует $\mu_{\boldsymbol{A}}(x)\boldsymbol{E}_n = (x\boldsymbol{E}_n - \boldsymbol{A})\Psi(x\boldsymbol{E}_n, \boldsymbol{A})$. Сравнивая полученное равенство с (1), заключаем, что

$$\boldsymbol{C}(x) = \Psi(x\boldsymbol{E}_n, \boldsymbol{A}). \tag{2}$$

Следующее утверждение дает критерий диагонализируемости матрицы.

Теорема 35. *Матрица $\boldsymbol{A} \in F^{n\times n}$ диагонализируема тогда и только тогда, когда ее минимальный многочлен $\mu_{\boldsymbol{A}}(x)$ разлагается на линейные множители над полем F и не имеет кратных корней.*

Доказательство. Пусть матрица $\boldsymbol{A}$ диагонализируема, т. е. для некоторой обратимой матрицы $\boldsymbol{T} \in F^{n\times n}$ матрица $\boldsymbol{D} = \boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T}$ является диагональной. Обозначим через $\lambda_1, \dots, \lambda_k$ все различные собственные значения матрицы $\boldsymbol{A}$. Легко видеть, что $(\boldsymbol{D}-\lambda_1\boldsymbol{E}_n)\dots(\boldsymbol{D}-\lambda_k\boldsymbol{E}_n)=\boldsymbol{O}$, поэтому и $(\boldsymbol{A}-\lambda_1\boldsymbol{E}_n)\dots(\boldsymbol{A}-\lambda_k\boldsymbol{E}_n)=\boldsymbol{O}$. Таким образом, $\boldsymbol{A}$ аннулируется многочленом $(x-\lambda_1)\dots(x-\lambda_k)$. Поскольку все корни $\lambda_1, \dots, \lambda_k$ должны быть корнями $\mu_{\boldsymbol{A}}(x)$, заключаем, что $\mu_{\boldsymbol{A}}(x) = (x-\lambda_1)\dots(x-\lambda_k)$.

Докажем достаточность. Предположим, что $\mu_{\boldsymbol{A}}(x) = (x-\lambda_1)\dots(x-\lambda_k)$, где $\lambda_1, \dots, \lambda_k$ – различные элементы поля F. Тогда в силу следствия 9 характеристический многочлен $\chi_{\boldsymbol{A}}(x)$ разлагается на линейные множители и имеет корни $\lambda_1, \dots, \lambda_k$. Как известно из курса линейной алгебры, пространство столбцов F_n допускает разложение в прямую сумму корневых подпространств относительно линейного оператора, определяемого матрицей $\boldsymbol{A}$. Покажем, что корневое подпространство V_j, соответствующее значению λ_j, есть ядро оператора $\boldsymbol{A} - \lambda_j\boldsymbol{E}_n$. Это будет следовать из того, что ядро $(\boldsymbol{A} - \lambda_j\boldsymbol{E}_n)^2$ совпадает с ядром $\boldsymbol{A} - \lambda_j\boldsymbol{E}_n$.

От противного, пусть $(\boldsymbol{A} - \lambda_j\boldsymbol{E}_n)^2 y = \boldsymbol{o}$, но $(\boldsymbol{A} - \lambda_j\boldsymbol{E}_n)y \neq \boldsymbol{o}$ для некоторого $y \in F_n$. Положим $z = (\boldsymbol{A} - \lambda_j\boldsymbol{E}_n)y$. Тогда $\boldsymbol{A}z = \lambda_j z$ и $(\boldsymbol{A} - \lambda_m\boldsymbol{E}_n)z = (\lambda_j - \lambda_m)z \neq \boldsymbol{o}$ при $m \neq j$. Имеем $\mu_{\boldsymbol{A}}(\boldsymbol{A})y =$

$= (\boldsymbol{A}-\lambda_1\boldsymbol{E}_n)\dots(\boldsymbol{A}-\lambda_k\boldsymbol{E}_n)y = \left(\prod_{m\neq j}(\boldsymbol{A}-\lambda_m\boldsymbol{E}_n)\right)(\boldsymbol{A}-\lambda_j\boldsymbol{E}_n)y =$
$= \left(\prod_{m\neq j}(\boldsymbol{A}-\lambda_m\boldsymbol{E}_n)\right) z = \prod_{m\neq j}(\lambda_j-\lambda_m)z \neq \boldsymbol{o}$, что невозможно.

Таким образом, пространство столбцов F_n есть прямая сумма ядер операторов $\boldsymbol{A}-\lambda_j\boldsymbol{E}_n$ при $j = 1, \dots, k$. Значит, существует базис из собственных векторов для линейного оператора, определяемого матрицей $\boldsymbol{A}$, и поэтому матрица $\boldsymbol{A}$ диагонализируема. Теорема доказана.

Следующее утверждение характеризует минимальный многочлен жордановой нормальной формы.

Предложение 10. *Если характеристический многочлен матрицы $\boldsymbol{A}$ порядка n разлагается на линейные множители, то ее минимальный многочлен равен $\prod_{j=1}^{k}(x-\lambda_j)^{s_j}$, где $\lambda_1, \dots, \lambda_k$ – различные собственные значения матрицы $\boldsymbol{A}$, s_j – максимальная степень элементарного делителя вида $(x-\lambda_j)^s$ x-матрицы $x\boldsymbol{E}_n-\boldsymbol{A}$.*

Доказательство. Ясно, что минимальные многочлены матрицы $\boldsymbol{A}$ и ее жордановой нормальной формы $\boldsymbol{J}$ совпадают. Легко видеть, что минимальный многочлен квазидиагональной матрицы равен общему наименьшему кратному минимальных многочленов ее диагональных клеток. Минимальный многочлен клетки Жордана порядка s, относящейся к собственному значению λ, совпадает с ее характеристическим многочленом и равен $(x-\lambda)^s$. Каждому элементарному делителю вида $(x-\lambda_j)^s$ x-матрицы $x\boldsymbol{E}_n-\boldsymbol{A}$ в жордановой нормальной форме $\boldsymbol{J}$ соответствует клетка Жордана порядка s, относящаяся к собственному значению λ_j. Из этого следует утверждение предложения.

Задачи для самостоятельного решения

1. Используя метод Д. К. Фаддеева, найти характеристический многочлен матрицы:

а) $\begin{pmatrix} -3 & -9 & -15 & -28 \\ 0 & -1 & -1 & -2 \\ -3 & -9 & -15 & -26 \\ 3 & 9 & 15 & 27 \end{pmatrix}$; б) $\begin{pmatrix} -1 & -4 & -8 & -11 & -4 & -8 \\ 0 & 1 & -1 & -1 & -1 & -1 \\ 0 & 0 & -1 & -1 & -1 & -1 \\ 1 & 1 & 3 & 6 & 1 & 2 \\ 1 & -1 & 0 & 1 & 1 & -1 \\ -1 & -1 & 0 & -1 & 2 & 4 \end{pmatrix}$.

2. Найти минимальный многочлен матрицы:

а) $\begin{pmatrix} 1 & 4 & 4 & 4 & 3 & 5 \\ 1 & 3 & 4 & 5 & 3 & 4 \\ 0 & -1 & 0 & -1 & -2 & -2 \\ 0 & -1 & -1 & 0 & 0 & -1 \\ 1 & 2 & 4 & 5 & 4 & 5 \\ -1 & -1 & -3 & -4 & -2 & -2 \end{pmatrix}$;

б) $\begin{pmatrix} 0 & 2 & 0 & -3 & 0 & 0 \\ -1 & 1 & -2 & -4 & -2 & -3 \\ 1 & 0 & 3 & 3 & 0 & 1 \\ 0 & -1 & -1 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 2 & 1 \\ 0 & 1 & 1 & 1 & 1 & 3 \end{pmatrix}$.

Глава 6

Прямое произведение матриц

§ 6.1. Определение и свойства прямого произведения матриц

Пусть F – поле, $\boldsymbol{A} \in F^{m\times m}$, $\boldsymbol{B} \in F^{n\times n}$. *Прямым произведением матрицы* $\boldsymbol{A}$ *на матрицу* $\boldsymbol{B}$ называется блочная матрица размеров $mn \times mn$:

$$\begin{pmatrix} a_{11}\boldsymbol{B} & a_{12}\boldsymbol{B} & \dots & a_{1m}\boldsymbol{B} \\ a_{21}\boldsymbol{B} & a_{22}\boldsymbol{B} & \dots & a_{2m}\boldsymbol{B} \\ \dots & \dots & \dots & \dots \\ a_{m1}\boldsymbol{B} & a_{m2}\boldsymbol{B} & \dots & a_{mm}\boldsymbol{B} \end{pmatrix},$$

которая обозначается через $\boldsymbol{A} \otimes \boldsymbol{B}$.

Для индексов $r, i = 1, \dots, m$, $s, j = 1, \dots, n$ имеем

$$(\boldsymbol{A} \otimes \boldsymbol{B})[(r-1)n + s, (i-1)n + j] = a_{ri}b_{sj}.$$

Свойства прямого произведения матриц собраны в следующем утверждении.

Теорема 36. *Имеют место следующие утверждения:*

1. Для любого $\mu \in F$ $(\mu\boldsymbol{A}) \otimes \boldsymbol{B} = \mu(\boldsymbol{A} \otimes \boldsymbol{B}) = \boldsymbol{A} \otimes (\mu\boldsymbol{B})$.

2. $(\boldsymbol{A} + \boldsymbol{B}) \otimes \boldsymbol{C} = \boldsymbol{A} \otimes \boldsymbol{C} + \boldsymbol{B} \otimes \boldsymbol{C}$.

3. $\boldsymbol{A} \otimes (\boldsymbol{B} + \boldsymbol{C}) = \boldsymbol{A} \otimes \boldsymbol{B} + \boldsymbol{A} \otimes \boldsymbol{C}$.
4. $\boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C}) = (\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C}$.
5. $(\boldsymbol{A} \otimes \boldsymbol{B})^\top = \boldsymbol{A}^\top \otimes \boldsymbol{B}^\top$.
6. Для матриц $\boldsymbol{A}, \boldsymbol{C} \in F^{m\times m}$, $\boldsymbol{B}, \boldsymbol{D} \in F^{n\times n}$ *справедливо*

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D}) = (\boldsymbol{AC}) \otimes (\boldsymbol{BD}).$$

7. $\boldsymbol{A} \otimes \boldsymbol{B} = (\boldsymbol{A} \otimes \boldsymbol{E}_n)(\boldsymbol{E}_m \otimes \boldsymbol{B})$.
8. $|\boldsymbol{A} \otimes \boldsymbol{B}| = |\boldsymbol{A}|^n |\boldsymbol{B}|^m$.
9. Если $|\boldsymbol{A}| \neq 0$ *и* $|\boldsymbol{B}| \neq 0$, *то* $(\boldsymbol{A} \otimes \boldsymbol{B})^{-1} = \boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1}$.
10. Если $\boldsymbol{A}_1, \ldots, \boldsymbol{A}_k \in F^{m\times m}$, $\boldsymbol{B}_1, \ldots, \boldsymbol{B}_k \in F^{n\times n}$, *то*

$$(\boldsymbol{A}_1 \otimes \boldsymbol{B}_1) \ldots (\boldsymbol{A}_k \otimes \boldsymbol{B}_k) = (\boldsymbol{A}_1 \ldots \boldsymbol{A}_k) \otimes (\boldsymbol{B}_1 \ldots \boldsymbol{B}_k).$$

11. $\mathsf{tr}(\boldsymbol{A} \otimes \boldsymbol{B}) = \mathsf{tr}(\boldsymbol{A})\mathsf{tr}(\boldsymbol{B})$.
12. $\mathsf{r}(\boldsymbol{A} \otimes \boldsymbol{B}) = \mathsf{r}(\boldsymbol{A})\mathsf{r}(\boldsymbol{B})$.

Доказательство. Утверждения 1–3 непосредственно следуют из определения линейных операций над матрицами.

Для доказательства утверждения 4 предположим, что $\boldsymbol{C} \in F^{p\times p}$ и $r, i = 1, \ldots, m$, $s, j = 1, \ldots, n$, $t, \ell = 1, \ldots, p$. Матрицы $(\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C}$ и $\boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C})$ имеют одинаковый порядок mnp. Имеем

$$(\boldsymbol{A} \otimes \boldsymbol{B})[(r-1)n + s, (i-1)n + j] = a_{ri} b_{sj},$$

$$(\boldsymbol{A} \otimes \boldsymbol{B}) \otimes \boldsymbol{C}[((r-1)n + s - 1)p + t, ((i-1)n + j - 1)p + \ell] = a_{ri} b_{sj} c_{t\ell}.$$

Далее, $(\boldsymbol{B} \otimes \boldsymbol{C})[(s-1)p + t, (j-1)p + \ell] = b_{sj} c_{t\ell}$ и

$$\boldsymbol{A} \otimes (\boldsymbol{B} \otimes \boldsymbol{C})([(r-1)np + (s-1)p + t, (i-1)np + (j-1)p + \ell] = a_{ri} b_{sj} c_{t\ell},$$

откуда следует требуемое.

Аналогично доказывается утверждение 5. Имеем

$$(\boldsymbol{A} \otimes \boldsymbol{B})^\top[(r-1)n + s, (i-1)n + j] =$$
$$= (\boldsymbol{A} \otimes \boldsymbol{B})[(i-1)n + j, (r-1)n + s] = a_{ir} b_{js}$$

и

$$(\boldsymbol{A}^\top \otimes \boldsymbol{B}^\top)[(r-1)n + s, (i-1)n + j] =$$
$$= \boldsymbol{A}^\top[r, i]\boldsymbol{B}^\top[s, j] = a_{ir} b_{js},$$

что и требуется доказать.

Докажем утверждение 6. Предположим, что $r, i, t = 1, \ldots, m$, $s, j, \ell = 1, \ldots, n$. Тогда

$$(\boldsymbol{A} \otimes \boldsymbol{B})[(r-1)n + s, (i-1)n + j] = a_{ri} b_{sj}$$

и

$$(\boldsymbol{C} \otimes \boldsymbol{D})[(i-1)n + j, (t-1)n + \ell] = c_{it} d_{j\ell}.$$

Далее,

$$\begin{gathered}
(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{C} \otimes \boldsymbol{D})[(r-1)n + s, (t-1)n + \ell] = \\
= \sum_{i=1}^{m} \sum_{j=1}^{n} (\boldsymbol{A} \otimes \boldsymbol{B})[(r-1)n + s, (i-1)n + j] \times \\
\times (\boldsymbol{C} \otimes \boldsymbol{D})[(i-1)n + j, (t-1)n + \ell] = \\
= \sum_{i=1}^{m} \sum_{j=1}^{n} a_{ri} b_{sj} c_{it} d_{j\ell} = (\sum_{i=1}^{m} a_{ri} c_{it})(\sum_{j=1}^{n} b_{sj} d_{j\ell}) = \\
= \boldsymbol{AC}[r, t] \boldsymbol{BD}[s, \ell] = (\boldsymbol{AC}) \otimes (\boldsymbol{BD})[(r-1)n + s, (t-1)n + \ell],
\end{gathered}$$

откуда следует требуемое.

Утверждение 7 непосредственно следует из утверждения 6.

Утверждение 8 следует из предыдущего, так как определитель $|\boldsymbol{A} \otimes \boldsymbol{E}_n|$ равен определителю квазитреугольной матрицы, у которой на главной диагонали n раз повторяется матрица $\boldsymbol{A}$ (которая получается четным числом перестановок строк и столбцов из исходной матрицы), и аналогично матрица $\boldsymbol{E}_m \otimes \boldsymbol{B}$ является квазитреугольной матрицей, у которой на главной диагонали m раз повторяется матрица $\boldsymbol{B}$.

Утверждение 9 следует из того, что, согласно утверждению 6,

$$(\boldsymbol{A} \otimes \boldsymbol{B})(\boldsymbol{A}^{-1} \otimes \boldsymbol{B}^{-1}) = \boldsymbol{E}_m \otimes \boldsymbol{E}_n = \boldsymbol{E}_{mn}.$$

Утверждение 10 следует по индукции из утверждения 6.

Утверждение 11 очевидно из определения.

Для доказательства утверждения 12 представим матрицу $\boldsymbol{A}$ в виде $\boldsymbol{A} = \boldsymbol{S}\boldsymbol{A}_1$, где $\boldsymbol{S}$ – невырожденная матрица, а $\boldsymbol{A}_1$ – ступенчатая по строкам матрица. Запишем $\boldsymbol{A} \otimes \boldsymbol{B} = (\boldsymbol{S} \otimes \boldsymbol{E}_n)(\boldsymbol{A}_1 \otimes \boldsymbol{B})$. Матрица $\boldsymbol{S} \otimes \boldsymbol{E}_n$ невырожденная, а $\mathsf{r}(\boldsymbol{A}_1 \otimes \boldsymbol{B}) = \mathsf{r}(\boldsymbol{A}_1)\mathsf{r}(\boldsymbol{B}) = \mathsf{r}(\boldsymbol{A})\mathsf{r}(\boldsymbol{B})$.

Теорема доказана.

§ 6.2. Составные матрицы

Пусть F – поле, $\varphi(x,y) \in F[x,y]$ – многочлен от двух переменных. Запишем $\varphi(x,y) = \sum_{i,j=0}^{p} \gamma_{ij} x^i y^j$ для некоторого натурального числа p и $\gamma_{ij} \in F$. Для матриц $\boldsymbol{A} \in F^{m\times m}$, $\boldsymbol{B} \in F^{n\times n}$ положим

$$\varphi(\boldsymbol{A},\boldsymbol{B}) = \sum_{i,j=0}^{p} \gamma_{ij} \boldsymbol{A}^i \otimes \boldsymbol{B}^j.$$

Составной матрицей называется матрица вида $\varphi(\boldsymbol{A},\boldsymbol{B})$.

Например, если $\varphi(x,y) = 2x + x^2y - xy^3$, то $\varphi(x,y) = 2x^1y^0 + {}+x^2y^1 - x^1y^3$ и $\varphi(\boldsymbol{A},\boldsymbol{B}) = 2\boldsymbol{A}\otimes\boldsymbol{E}_n + \boldsymbol{A}^2\otimes\boldsymbol{B} - \boldsymbol{A}\otimes\boldsymbol{B}^3$.

Одна из главных причин интереса к прямому произведению матриц – простая связь между собственными значениями исходных матриц $\boldsymbol{A}$, $\boldsymbol{B}$ и составной матрицы $\varphi(\boldsymbol{A},\boldsymbol{B})$.

Теорема 37. *Пусть $\boldsymbol{A}$, $\boldsymbol{B}$ – матрицы, у которых характеристический многочлен разлагается на линейные множители, и $\lambda_1,\dots,\lambda_m$ – все собственные значения матрицы $\boldsymbol{A}$, $\mu_1,\dots,\mu_n$ – все собственные значения матрицы $\boldsymbol{B}$. Тогда собственные значения матрицы $\varphi(\boldsymbol{A},\boldsymbol{B})$ исчерпываются mn скалярами $\varphi(\lambda_i,\mu_j)$ при всевозможных $i = 1,\dots,m$, $j = 1,\dots,n$.*

Доказательство. Приведем матрицы $\boldsymbol{A}$, $\boldsymbol{B}$ к жордановой нормальной форме. Пусть $\boldsymbol{S} \in F^{m\times m}$, $\boldsymbol{T} \in F^{n\times n}$ – невырожденные матрицы и $\boldsymbol{S}^{-1}\boldsymbol{A}\boldsymbol{S} = \boldsymbol{J}_1$, $\boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T} = \boldsymbol{J}_2$ – жордановы нормальные формы. Так как $\boldsymbol{J}_1$ – верхнетреугольная матрица с элементами $\lambda_1,\dots,\lambda_m$ на главной диагонали, ее любая степень $\boldsymbol{J}_1^i$ – также верхнетреугольная матрица с элементами $\lambda_1^i,\dots,\lambda_m^i$ на главной диагонали. Аналогично любая степень $\boldsymbol{J}_2^j$ есть верхнетреугольная матрица с элементами $\mu_1^j,\dots,\mu_n^j$ на главной диагонали. Легко видеть, что прямое произведение $\boldsymbol{J}_1^i\otimes\boldsymbol{J}_2^j$ есть верхнетреугольная матрица с mn элементами вида $\lambda_s^i\mu_t^j$ на главной диагонали. Следовательно, $\varphi(\boldsymbol{J}_1,\boldsymbol{J}_2)$ есть верхнетреугольная матрица, на главной диагонали которой находится mn скаляров вида $\varphi(\lambda_s,\mu_t)$ ($s = 1,\dots,m$, $t = 1,\dots,n$).

Покажем, что матрица $\varphi(\boldsymbol{A},\boldsymbol{B})$ подобна матрице $\varphi(\boldsymbol{J}_1,\boldsymbol{J}_2)$. Имеем

$$\boldsymbol{J}_1^i\otimes\boldsymbol{J}_2^j = (\boldsymbol{S}^{-1}\boldsymbol{A}^i\boldsymbol{S})\otimes(\boldsymbol{T}^{-1}\boldsymbol{B}^j\boldsymbol{T}) =$$

$$= (\boldsymbol{S}^{-1} \otimes \boldsymbol{T}^{-1})(\boldsymbol{A}^i \otimes \boldsymbol{B}^j)(\boldsymbol{S} \otimes \boldsymbol{T}) = (\boldsymbol{S} \otimes \boldsymbol{T})^{-1}(\boldsymbol{A}^i \otimes \boldsymbol{B}^j)(\boldsymbol{S} \otimes \boldsymbol{T}).$$

Мы видим, что матрицы $\boldsymbol{A}^i \otimes \boldsymbol{B}^j$ и $\boldsymbol{J}_1^i \otimes \boldsymbol{J}_2^j$ подобны. Следовательно, матрицы $\varphi(\boldsymbol{A}, \boldsymbol{B})$ и $\varphi(\boldsymbol{J}_1, \boldsymbol{J}_2)$ подобны и потому имеют одинаковые собственные значения. Теорема доказана.

Отметим важные частные случаи, получаемые с помощью теоремы 37.

Следствие 10. *Пусть $\boldsymbol{A}$, $\boldsymbol{B}$ – матрицы, у которых характеристический многочлен разлагается на линейные множители, и $\lambda_1, \ldots, \lambda_m$ – все собственные значения матрицы $\boldsymbol{A}$, $\mu_1, \ldots, \mu_n$ – все собственные значения матрицы $\boldsymbol{B}$. Тогда собственные значения матрицы $\boldsymbol{A} \otimes \boldsymbol{B}$ исчерпываются mn произведениями $\lambda_i \mu_j$ при всевозможных $i = 1, \ldots, m$, $j = 1, \ldots, n$.*

Следствие 11. *Пусть $\boldsymbol{A}$, $\boldsymbol{B}$ – матрицы, у которых характеристический многочлен разлагается на линейные множители, и $\lambda_1, \ldots, \lambda_m$ – все собственные значения матрицы $\boldsymbol{A}$, $\mu_1, \ldots, \mu_n$ – все собственные значения матрицы $\boldsymbol{B}$. Тогда собственные значения матрицы $\boldsymbol{A} \otimes \boldsymbol{E}_n + \boldsymbol{E}_m \otimes \boldsymbol{B}$ исчерпываются mn суммами $\lambda_i + \mu_j$ при всевозможных $i = 1, \ldots, m$, $j = 1, \ldots, n$.*

Собственные векторы прямого произведения матриц описывает следующее

Предложение 11. *Пусть $\boldsymbol{A} \in F^{m \times m}$, $\boldsymbol{B} \in F^{n \times n}$, λ – собственное значение матрицы $\boldsymbol{A}$, $\boldsymbol{x} = (\xi_1, \ldots, \xi_m)^\top$ – соответствующий собственный вектор и μ – собственное значение матрицы $\boldsymbol{B}$, $\boldsymbol{y} = (\nu_1, \ldots, \nu_n)^\top$ – соответствующий собственный вектор. Тогда $\lambda\mu$ – собственное значение матрицы $\boldsymbol{A} \otimes \boldsymbol{B}$ и $\boldsymbol{z} = (\xi_1\nu_1, \ldots, \xi_m\nu_1, \ldots, \xi_1\nu_n, \ldots, \xi_m\nu_n)^\top$ – соответствующий собственный вектор.*

Доказательство. Пусть $\boldsymbol{X} = \begin{pmatrix} \xi_1 & 0 & 0 & 0 \\ \xi_2 & 0 & 0 & 0 \\ \cdots & \cdots & \cdots & \cdots \\ \xi_m & 0 & 0 & 0 \end{pmatrix} \in F^{m \times m}$,

$$\boldsymbol{Y} = \begin{pmatrix} \nu_1 & 0 & 0 & 0 \\ \nu_2 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ \nu_n & 0 & 0 & 0 \end{pmatrix} \in F^{n\times n} \text{ и}$$

$$\boldsymbol{Z} = \boldsymbol{X} \otimes \boldsymbol{Y} = \begin{pmatrix} \xi_1\nu_1 & 0 & 0 & 0 \\ \xi_1\nu_2 & 0 & 0 & 0 \\ \dots & \dots & \dots & \dots \\ \xi_m\nu_n & 0 & 0 & 0 \end{pmatrix} \in F^{mn\times mn}.$$

Имеем $\boldsymbol{AX} = \lambda\boldsymbol{X}$, $\boldsymbol{BY} = \mu\boldsymbol{Y}$, $(\boldsymbol{A}\otimes\boldsymbol{B})\boldsymbol{Z} = (\boldsymbol{A}\otimes\boldsymbol{B})(\boldsymbol{X}\otimes\boldsymbol{Y}) = (\boldsymbol{AX})\otimes(\boldsymbol{BY}) = (\lambda\boldsymbol{X})\otimes(\mu\boldsymbol{Y}) = \lambda\mu\boldsymbol{Z}$. Следовательно,

$$(\boldsymbol{A}\otimes\boldsymbol{B})\boldsymbol{z} = \lambda\mu\boldsymbol{z}.$$

Предложение доказано.

Рассмотрим матричное уравнение вида

$$\sum_{j=1}^{k} \boldsymbol{A}_j\boldsymbol{X}\boldsymbol{B}_j = \boldsymbol{C}, \tag{1}$$

где $\boldsymbol{A}_j, \boldsymbol{B}_j, \boldsymbol{C} \in F^{n\times n}$ $(j = 1,\dots,k)$ – известные матрицы, $\boldsymbol{X} \in F^{n\times n}$ – неизвестная матрица. Сведем решение этого уравнения к решению уравнения $\boldsymbol{Gx} = \boldsymbol{c}$, где $\boldsymbol{G} \in F^{mn\times mn}$, $\boldsymbol{x}$ и $\boldsymbol{c}$ – столбцы всех элементов (по строкам) матриц $\boldsymbol{X}$ и $\boldsymbol{C}$ соответственно, т. е. $\boldsymbol{x} = (\boldsymbol{X}_{1\bullet},\dots,\boldsymbol{X}_{n\bullet})^\top$ и $\boldsymbol{c} = (\boldsymbol{C}_{1\bullet},\dots,\boldsymbol{C}_{n\bullet})^\top$.

Сначала рассмотрим случай одного слагаемого в правой части (1): $\boldsymbol{AXB} = \boldsymbol{C}$. Для индексов r, s имеем $\boldsymbol{AXB}[r,s] = \boldsymbol{C}[r,s]$ или

$$\sum_{i=1}^{n}\sum_{j=1}^{n} \boldsymbol{A}[r,i]\boldsymbol{X}[i,j]\boldsymbol{B}[j,s] = \boldsymbol{C}[r,s].$$

Последнее равенство запишем в виде

$$\sum_{i=1}^{n}\sum_{j=1}^{n} (\boldsymbol{A}[r,i]\boldsymbol{B}[j,s])\boldsymbol{X}[i,j] = \boldsymbol{C}[r,s].$$

При $r, s = 1,\dots,n$ получается система из n^2 уравнений относительно незвестных элементов матрицы $\boldsymbol{X}$. Уравнения будем нумеровать, придавая s значения $1,\dots,n$ при фиксированном r. Тогда в $((r-1)n+s)$-м уравнении коэффициент при $\boldsymbol{X}[i,j]$ равен

$\boldsymbol{A}[r,i]\boldsymbol{B}[j,s]$. Таким образом, матрица $\boldsymbol{G}$ в данном случае удовлетворяет условию $\boldsymbol{G}[(r-1)n+s,(i-1)n+j]=a_{ri}b_{js}$, т. е.

$$\boldsymbol{G}=\boldsymbol{A}\otimes\boldsymbol{B}^{\top}.$$

Таким образом, уравнение $\boldsymbol{AXB}=\boldsymbol{C}$ равносильно уравнению

$$(\boldsymbol{A}\otimes\boldsymbol{B}^{\top})\boldsymbol{x}=\boldsymbol{c}.$$

Возвращаясь к уравнению (1), заключаем, что оно равносильно уравнению

$$\sum_{j=1}^{k}(\boldsymbol{A}_j\otimes\boldsymbol{B}_j^{\top})\boldsymbol{x}=\boldsymbol{c}.$$

Предложение 12. *Пусть F – алгебраически замкнутое поле, $\boldsymbol{A}\in F^{n\times n}$, $\lambda\in F$. Уравнение $\boldsymbol{AX}-\boldsymbol{XA}=\lambda\boldsymbol{X}$ имеет ненулевое решение тогда и только тогда, когда $\lambda=\mu_i-\mu_j$ для некоторых собственных значений матрицы $\boldsymbol{A}$.*

Для доказательства заметим, что рассматриваемое уравнение равносильно уравнению $(\boldsymbol{A}\otimes\boldsymbol{E}_n-\boldsymbol{E}_n\otimes\boldsymbol{A}^{\top})\boldsymbol{x}=\lambda\boldsymbol{x}$. Так как собственные значения матриц $\boldsymbol{A}$ и $\boldsymbol{A}^{\top}$ одни и те же, собственные значения матриц $\boldsymbol{A}\otimes\boldsymbol{E}_n-\boldsymbol{E}_n\otimes\boldsymbol{A}^{\top}$ и $\boldsymbol{A}\otimes\boldsymbol{E}_n-\boldsymbol{E}_n\otimes\boldsymbol{A}$ одинаковы. Последняя матрица в качестве собственных значений имеет всевозможные разности собственных значений матрицы $\boldsymbol{A}$, откуда следует требуемое.

Предложение 13. *Пусть F – алгебраически замкнутое поле, $\boldsymbol{A},\boldsymbol{B},\boldsymbol{C}\in F^{n\times n}$. Уравнение $\boldsymbol{AX}+\boldsymbol{XB}=\boldsymbol{C}$ имеет единственное решение тогда и только тогда, когда матрицы $\boldsymbol{A}$ и $-\boldsymbol{B}$ не имеют общих собственных значений.*

Указанное уравнение равносильно уравнению

$$(\boldsymbol{A}\otimes\boldsymbol{E}_n+\boldsymbol{E}_n\otimes\boldsymbol{B}^{\top})\boldsymbol{x}=\boldsymbol{c}. \tag{2}$$

Как и в доказательстве предложения 1, замечаем, что собственные значения матриц $\boldsymbol{A}\otimes\boldsymbol{E}_n+\boldsymbol{E}_n\otimes\boldsymbol{B}^{\top}$ и $\boldsymbol{A}\otimes\boldsymbol{E}_n+\boldsymbol{E}_n\otimes\boldsymbol{B}$ одни и те же. Матричное уравнение (2) имеет единственное решение тогда и только тогда, когда матрица $\boldsymbol{A} \otimes \boldsymbol{E}_n + \boldsymbol{E}_n \otimes \boldsymbol{B}^\top$ является невырожденной, т. е. когда она не имеет нулевых собственных значений. Матрица $\boldsymbol{A} \otimes \boldsymbol{E}_n + \boldsymbol{E}_n \otimes \boldsymbol{B}$ не имеет нулевых собственных значений тогда и только тогда, когда матрицы $\boldsymbol{A}$ и $-\boldsymbol{B}$ не имеют общих собственных значений.

Задачи для самостоятельного решения

1. Решить матричное уравнение $\boldsymbol{A}_1\boldsymbol{X}\boldsymbol{B}_1 + \boldsymbol{A}_2\boldsymbol{X}\boldsymbol{B}_2 = \boldsymbol{C}$, где

а) $\boldsymbol{A}_1 = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$, $\boldsymbol{B}_1 = \begin{pmatrix} -1 & 3 \\ 4 & -5 \end{pmatrix}$, $\boldsymbol{A}_2 = \begin{pmatrix} -1 & -1 \\ 2 & -3 \end{pmatrix}$, $\boldsymbol{B}_2 = \begin{pmatrix} -2 & 3 \\ 8 & -1 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} 37 & -81 \\ 295 & -52 \end{pmatrix}$;

б) $\boldsymbol{A}_1 = \begin{pmatrix} 4 & -2 \\ -3 & 1 \end{pmatrix}$, $\boldsymbol{B}_1 = \begin{pmatrix} 5 & 3 \\ 4 & 1 \end{pmatrix}$, $\boldsymbol{A}_2 = \begin{pmatrix} 1 & 1 \\ -2 & 3 \end{pmatrix}$, $\boldsymbol{B}_2 = \begin{pmatrix} 2 & -3 \\ -8 & 1 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} 210 & 27 \\ -65 & 56 \end{pmatrix}$.

2. Определить, имеет ли матричное уравнение $\boldsymbol{A}\boldsymbol{X} + \boldsymbol{X}\boldsymbol{B} = \boldsymbol{C}$ единственное решение в следующих случаях:

а) $\boldsymbol{A} = \begin{pmatrix} 5 & 3 \\ 3 & 5 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 3 & 1 \\ 1 & 3 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} 8 & 1 \\ -1 & 7 \end{pmatrix}$;

б) $\boldsymbol{A} = \begin{pmatrix} 7 & 1 \\ 1 & 7 \end{pmatrix}$, $\boldsymbol{B} = \begin{pmatrix} 8 & 3 \\ 3 & 8 \end{pmatrix}$, $\boldsymbol{C} = \begin{pmatrix} 3 & 4 \\ -5 & 1 \end{pmatrix}$.

Глава 7

Функции от матриц

§ 7.1. Определение функции от матрицы

В этой главе рассматриваются матрицы над полем комплексных чисел.

Построим интерполяционный многочлен Эрмита. Пусть $\lambda_1, \ldots, \lambda_s$ – различные комплексные числа, $m_1, \ldots, m_s$ – некоторые натуральные числа, f_{jp} $(p = 0, \ldots, m_j - 1,\ j = 1, \ldots, s)$ – фиксированные комплексные числа. Требуется построить многочлен $f(x)$, удовлетворяющий условиям

$$f(\lambda_j) = f_{j0},\ f^{(p)}(\lambda_j) = f_{jp}\ (p = 1, \ldots, m_j - 1,\ j = 1, \ldots, s)$$

и имеющий степень $m = \left(\sum_{j=1}^{s} m_j\right) - 1$.

Сначала построим систему многочленов $\varphi_{jk}(x)$ $(k = 1, \ldots, m_j,$ $j = 1, \ldots, s)$, удовлетворяющих следующим условиям:

$$\varphi_{jk}(\lambda_\ell) = 0,\ \varphi_{jk}^{(p)}(\lambda_\ell) = 0$$

при $p = 1, \ldots, m_\ell - 1,\ \ell \neq k$, и

$$\varphi_{jk}^{(j-1)}(\lambda_k) = 1,\ \varphi_{jk}^{(p)}(\lambda_k) = 0$$

при $p \neq j-1,\ p = 0, 1, \ldots, m_j - 1$. Используем метод неопределенных коэффициентов. Пусть $\varphi_{jk}(x) = \alpha_0 + \alpha_1 x + \ldots + \alpha_m x^m$. Имеется m

коэффициентов и m линейных уравнений для них. Получающаяся система совместна, как известно из курса «Методы приближенных вычислений». Она имеет единственное решение, так как для любых двух решений $\varphi(x)$ и $\varphi^*(x)$ их разность $\varphi(x) - \varphi^*(x)$ имеет степень, не превосходящую m, и s различных корней кратностей в сумме больше m, т. е. является нулевым многочленом.

Построив многочлены $\varphi_{jk}(x)$, положим

$$f(x) = \sum_{j=1}^{s} \sum_{k=1}^{m_j} f_{j,k-1} \varphi_{jk}(x).$$

Полученный многочлен называется *интерполяционным многочленом Эрмита.*

Пусть $\boldsymbol{A} \in \mathbb{C}^{n \times n}$ и $\mu_{\boldsymbol{A}}(x) = (x - \lambda_1)^{m_1} \ldots (x - \lambda_s)^{m_s}$. Положим $m = m_1 + \ldots + m_s$.

Предложение 14. *Пусть $f(x), g(x) \in \mathbb{C}[x]$. Тогда равенство $f(\boldsymbol{A}) = g(\boldsymbol{A})$ имеет место тогда и только тогда, когда $f(x) - g(x)$ делится на $\mu_{\boldsymbol{A}}(x)$, что равносильно системе равенств*

$$f(\lambda_k) = g(\lambda_k),\ f^{(p)}(\lambda_k) = g^{(p)}(\lambda_k),\ p = 1, \ldots, m_k - 1,\ k = 1, \ldots, s.$$

Доказательство следует из основного свойства минимального многочлена: если $h(x) \in \mathbb{C}[x]$ и $h(\boldsymbol{A}) = 0$, то $\mu_{\boldsymbol{A}}(x)$ делит $h(x)$.

Пусть $\Psi(x)$ – функция, дифференцируемая в точках $\lambda_1, \ldots, \lambda_s$ достаточное число раз. Тогда *значением функции $\Psi(x)$ на спектре матрицы $\boldsymbol{A}$* называется набор чисел

$$\Psi(\lambda_k),\ \Psi^{(p)}(\lambda_k),\ p = 1, \ldots, m_k - 1,\ k = 1, \ldots, s.$$

В этом случае также говорят, что функция $\Psi(x)$ *определена на спектре* матрицы $\boldsymbol{A}$.

По значениям функции на спектре матрицы строится интерполяционный многочлен Эрмита

$$H(x) = \sum_{k=1}^{s} \sum_{p=1}^{m_k} \Psi^{(p-1)}(\lambda_k) \varphi_{kp}(x).$$

Для него имеют место равенства

$$\Psi(\lambda_k) = H(\lambda_k),\ \Psi^{(p)}(\lambda_k) = H^{(p)}(\lambda_k),\ p = 1, \ldots, m_k - 1,\ k = 1, \ldots, s.$$

Значением функции $\Psi(x)$ *от матрицы* $\boldsymbol{A}$ называется матрица $H(\boldsymbol{A})$. Обозначение: $\Psi(\boldsymbol{A})$. Так как интерполяционный многочлен Эрмита определяется однозначно по значениям функции на спектре матрицы, значение функции от матрицы определено однозначно.

Для любой функции, определенной на спектре матрицы $\boldsymbol{A}$, определено значение этой функции от матрицы $\boldsymbol{A}$.

Рассмотрим несколько примеров.

Пример 1. Вычислить $f(\boldsymbol{A})$, где $f(x) = e^{\lambda x}$, $\boldsymbol{A} = \begin{pmatrix} 6 & -1 \\ 3 & 2 \end{pmatrix}$.

Вычислим $\chi_{\boldsymbol{A}}(x) = \begin{vmatrix} x-6 & 1 \\ -3 & x-2 \end{vmatrix} = x^2 - 8x + 15 = (x-3)(x-5)$. Следовательно, $\mu_{\boldsymbol{A}}(x) = \chi_{\boldsymbol{A}}(x)$. Интерполяционный многочлен Эрмита имеет степень 1: $H(x) = \alpha + \beta x$. Имеем $H(3) = f(3) = e^{3\lambda}$, $H(5) = f(5) = e^{5\lambda}$, откуда $\begin{cases} \alpha + 3\beta = e^{3\lambda}, \\ \alpha + 5\beta = e^{5\lambda}. \end{cases}$ Решая систему, получаем $\beta = \frac{1}{2}(e^{5\lambda} - e^{3\lambda})$, $\alpha = \frac{1}{2}(5e^{3\lambda} - 3e^{5\lambda})$. Следовательно, $H(x) = \frac{1}{2}((5e^{3\lambda} - 3e^{5\lambda}) + (e^{5\lambda} - e^{3\lambda})x)$. Далее,

$$H(\boldsymbol{A}) = \frac{1}{2}((5e^{3\lambda} - 3e^{5\lambda})\boldsymbol{E}_2 + (e^{5\lambda} - e^{3\lambda})\boldsymbol{A}) =$$

$$= \frac{1}{2}\begin{pmatrix} 3e^{5\lambda} - e^{3\lambda} & e^{3\lambda} - e^{5\lambda} \\ 3e^{5\lambda} - 3e^{3\lambda} & 3e^{3\lambda} - e^{5\lambda} \end{pmatrix} =$$

$$= \frac{1}{2}e^{5\lambda}\begin{pmatrix} 3 & -1 \\ 3 & -1 \end{pmatrix} + \frac{1}{2}e^{3\lambda}\begin{pmatrix} -1 & -1 \\ -3 & 3 \end{pmatrix}.$$

Итак, $e^{\boldsymbol{A}} = \frac{1}{2}e^{5\lambda}\begin{pmatrix} 3 & -1 \\ 3 & -1 \end{pmatrix} + \frac{1}{2}e^{3\lambda}\begin{pmatrix} -1 & -1 \\ -3 & 3 \end{pmatrix}$.

Пример 2. Вычислить $f(\boldsymbol{A})$, где $f(x) = \sin x$, $\boldsymbol{A} = \frac{\pi}{2}\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix}$.

Вычислим $\mu_{\boldsymbol{A}}(x) = \chi_{\boldsymbol{A}}(x) = \begin{vmatrix} x-\pi & 0 & 0 \\ 0 & x-\pi/2 & -1 \\ 0 & 0 & x-\pi/2 \end{vmatrix} = (x-\pi)(x-\pi/2)^2$. Интерполяционный многочлен Эрмита имеет степень 2: $H(x) = \alpha + \beta x + \gamma x^2$. Имеем $H(\pi) = f(\pi) = \sin\pi = 0$, $H(\pi/2) = f(\pi/2) = \sin\pi/2 = 1$, $H'(\pi/2) = f'(\pi/2) = \cos\pi/2 = 0$, откуда $\begin{cases} \alpha + \pi\beta + \pi^2\gamma = 0, \\ \alpha + \frac{\pi}{2}\beta + \frac{\pi^2}{4}\gamma = 1, \\ \beta + \pi\gamma = 0. \end{cases}$

Решая систему, получаем $\alpha = 0$, $\beta = -\frac{4}{\pi}$, $\gamma = \frac{4}{\pi^2}$. Следовательно, $H(x) = -\frac{4}{\pi}x + \frac{4}{\pi^2}x^2$. Тогда $\sin \boldsymbol{A} = -\frac{4}{\pi}\boldsymbol{A} + \frac{4}{\pi^2}\boldsymbol{A}^2 =$
$= -2\begin{pmatrix} 2 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{pmatrix} + \begin{pmatrix} 4 & 0 & 0 \\ 0 & 1 & 2 \\ 0 & 0 & 1 \end{pmatrix} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & -1 & 0 \\ 0 & 0 & -1 \end{pmatrix}$.

Пример 3. Найти $f(\boldsymbol{J})$ для матрицы $\boldsymbol{J} = \boldsymbol{J}_n(0) =$

$= \begin{pmatrix} 0 & 1 & 0 & \cdots & 0 \\ 0 & 0 & 1 & \cdots & 0 \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & 1 \\ 0 & 0 & 0 & \cdots & 0 \end{pmatrix}$ и функции $f(x)$, имеющей производные до $n-1$-го порядка включительно в точке 0.

Минимальный многочлен $\mu_{\boldsymbol{J}}(x) = x^n$. Значения $f(x)$ на спектре $\boldsymbol{J}$ суть $f(0)$, $f'(0)$, $f''(0), \ldots, f^{(n-1)}(0)$. Тогда $H(x) =$

$= f(0) + f'(0)x + \frac{1}{2}f''(0)x^2 + \ldots + \frac{1}{(n-1)!}f^{(n-1)}(0)x^{n-1}$ и $f(\boldsymbol{J}) =$
$= f(0)\boldsymbol{E}_n + f'(0)\boldsymbol{J} + \frac{1}{2}f''(0)\boldsymbol{J}^2 + \ldots + \frac{1}{(n-1)!}f^{(n-1)}(0)\boldsymbol{J}^{n-1} =$

$= \begin{pmatrix} f(0) & f'(0) & \frac{1}{2}f''(0) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(0) \\ 0 & f(0) & f'(0) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(0) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & f'(0) \\ 0 & 0 & 0 & \cdots & f(0) \end{pmatrix}$, так как при возведении матрицы $\boldsymbol{J}$ в степень $1 < k < n$ диагональ из единиц сдвигается на k позиций вправо.

Пример 4. Для матрицы $\boldsymbol{J} = \boldsymbol{J}_n(\lambda_0) = \lambda_0\boldsymbol{E}_n + \boldsymbol{J}_n(0)$ и функции $f(x)$, имеющей производные до $n-1$-го порядка включительно в точке λ_0, найти $f(\boldsymbol{J})$.

Минимальный многочлен $\mu_{\boldsymbol{J}}(x) = (x - \lambda_0)^n$. Значения $f(x)$ на спектре $\boldsymbol{J}$ суть $f(\lambda_0)$, $f'(\lambda_0)$, $f''(\lambda_0), \ldots, f^{(n-1)}(\lambda_0)$. Тогда

$$H(x) = f(\lambda_0) + f'(0)(x-\lambda_0) + \frac{1}{2}f''(\lambda_0)(x-\lambda_0)^2 + \ldots + \\ + \frac{1}{(n-1)!}f^{(n-1)}(\lambda_0)(x-\lambda_0)^{n-1}$$

и

$$f(\boldsymbol{J}) = f(\lambda_0)\boldsymbol{E}_n + f'(\lambda_0)\boldsymbol{J}_n(0) + \frac{1}{2}f''(\lambda_0)\boldsymbol{J}_n(0)^2 + \ldots + \\ + \frac{1}{(n-1)!}f^{(n-1)}(\lambda_0)\boldsymbol{J}_n(0)^{n-1} = \\ = \begin{pmatrix} f(\lambda_0) & f'(\lambda_0) & \frac{1}{2}f''(\lambda_0) & \cdots & \frac{1}{(n-1)!}f^{(n-1)}(\lambda_0) \\ 0 & f(\lambda_0) & f'(\lambda_0) & \cdots & \frac{1}{(n-2)!}f^{(n-2)}(\lambda_0) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ 0 & 0 & 0 & \cdots & f'(\lambda_0) \\ 0 & 0 & 0 & \cdots & f(\lambda_0) \end{pmatrix}.$$

Теорема 38. *Пусть $\boldsymbol{A} = [\boldsymbol{A}_1, \ldots, \boldsymbol{A}_p]$ – квазидиагональная матрица и функция $f(x)$ определена на спектре матрицы $\boldsymbol{A}$. Тогда матрица $f(\boldsymbol{A}) = [f(\boldsymbol{A}_1), \ldots, f(\boldsymbol{A}_p)]$ является квазидиагональной и на ее главной диагонали расположены блоки – значения функции $f(x)$ от соответствующих диагональных блоков матрицы $\boldsymbol{A}$.*

Доказательство. Легко видеть, что для любого многочлена $g(x) \in \mathbb{C}[x]$ справедливо равенство $g(\boldsymbol{A}) = [g(\boldsymbol{A}_1), \ldots, g(\boldsymbol{A}_p)]$. Пусть $H(x)$ – интерполяционный многочлен Эрмита для функции $f(x)$ на спектре матрицы $\boldsymbol{A}$. Тогда $f(\boldsymbol{A}) = H(\boldsymbol{A}) = [H(\boldsymbol{A}_1), \ldots, H(\boldsymbol{A}_p)]$. Ясно, что $\mu_{\boldsymbol{A}}(x)$ есть наименьшее общее кратное многочленов $\mu_{\boldsymbol{A}_j}(x)$ при $j = 1, \ldots, p$. Поэтому из того, что функция $f(x)$ определена на спектре матрицы $\boldsymbol{A}$, следует, что она определена и на спектре любой матрицы $\boldsymbol{A}_j$, $j = 1, \ldots, p$. Отсюда следует, что значения $f(x)$ и $H(x)$ совпадают на спектре любой матрицы $\boldsymbol{A}_j$ и, значит, $f(\boldsymbol{A}_j) = H(\boldsymbol{A}_j)$. Следовательно, $f(\boldsymbol{A}) = [f(\boldsymbol{A}_1), \ldots, f(\boldsymbol{A}_p)]$.

Теорема 39. *Пусть $\boldsymbol{A}, \boldsymbol{B} \in \mathbb{C}^{n\times n}$, $\boldsymbol{A} = \boldsymbol{T}^{-1}\boldsymbol{B}\boldsymbol{T}$, и функция $f(x)$ определена на спектре матрицы $\boldsymbol{B}$. Тогда функция $f(x)$ определена на спектре матрицы $\boldsymbol{A}$ и $f(\boldsymbol{A}) = \boldsymbol{T}^{-1}f(\boldsymbol{B})\boldsymbol{T}$.*

Доказательство. Так как $\mu_{\boldsymbol{A}}(x) = \mu_{\boldsymbol{B}}(x)$, функция $f(x)$ определена на спектре матрицы $\boldsymbol{A}$. Поэтому интерполяционный многочлен $H(x)$ для матриц $\boldsymbol{A}, \boldsymbol{B}$ получается один и тот же. Отсюда следует, что $f(\boldsymbol{A}) = H(\boldsymbol{A})$ и $f(\boldsymbol{B}) = H(\boldsymbol{B})$. Так как $H(\boldsymbol{A}) = \boldsymbol{T}^{-1}H(\boldsymbol{B})\boldsymbol{T}$, получаем требуемое.

Пусть матрица $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ и функция $f(x)$ определена на спектре матрицы $\boldsymbol{A}$.

Если матрица $\boldsymbol{A}$ диагонализируема, т. е. $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = [\lambda_1, \ldots, \lambda_n]$, то $f(\boldsymbol{A}) = \boldsymbol{T}[f(\lambda_1), \ldots, f(\lambda_n)]\boldsymbol{T}^{-1}$.

Если матрица $\boldsymbol{A}$ приводится к жордановой нормальной форме $\boldsymbol{T}^{-1}\boldsymbol{A}\boldsymbol{T} = [\boldsymbol{J}_{m_1}(\lambda_1), \ldots, \boldsymbol{J}_{m_p}(\lambda_p)]$, то $f(\boldsymbol{A}) = \boldsymbol{T}[f(\boldsymbol{J}_{m_1}(\lambda_1)), \ldots, f(\boldsymbol{J}_{m_p}(\lambda_p))]\boldsymbol{T}^{-1}$.

Предложение 15. *Если $\lambda_1, \ldots, \lambda_n$ – все собственные значения (с учетом кратности) матрицы $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ и функция $f(x)$ определена на спектре матрицы $\boldsymbol{A}$, то собственные значения матрицы $f(\boldsymbol{A})$ исчерпываются числами $f(\lambda_1), \ldots, f(\lambda_n)$.*

Доказательство. Если матрица $\boldsymbol{A}$ диагонализируема, то требуемое непосредственно следует из сказанного выше. Если матрица $\boldsymbol{A}$ приводится к жордановой нормальной форме, то собственные значения матрицы $f(\boldsymbol{A})$ исчерпываются собственными значениями клеток $f(\boldsymbol{J}_{m_1}(\lambda_1)), \ldots, f(\boldsymbol{J}_{m_p}(\lambda_p))$ на главной диагонали. Вспоминая пример 4, заключаем, что $f(\lambda_j)$ – единственное собственное значение матрицы $f(\boldsymbol{J}_{m_j}(\lambda_j))$.

§ 7.2. Компоненты матрицы

Теорема 40. *Пусть матрица $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, функция $f(x)$ определена на спектре матрицы $\boldsymbol{A}$, $\mu_{\boldsymbol{A}}(x) = (x-\lambda_1)^{m_1}\ldots(x-\lambda_s)^{m_s}$, $f_k^{(j)} = f^{(j)}(\lambda_k)$ ($j = 0, \ldots, m_k - 1, k = 1, \ldots, s$ и $f_k^{(0)} = f(\lambda_k)$). Тогда существуют такие независимые от f матрицы $\boldsymbol{Z}_{kj}$, что $f(\boldsymbol{A}) = \sum_{k=1}^{s}\sum_{j=1}^{m_j} f_k^{(j-1)}\boldsymbol{Z}_{kj}$. Матрицы $\boldsymbol{Z}_{kj}$ линейно независимы и перестановочны с матрицей $\boldsymbol{A}$.*

Доказательство. Вспомним представление интерполяционного многочлена Эрмита для функции $f(x)$ на спектре матрицы

$\boldsymbol{A}$: $H(x) = \sum_{j=1}^{s} \sum_{k=1}^{m_j} f_k^{(j-1)} \varphi_{jk}(x)$. Отсюда

$$f(\boldsymbol{A}) = H(\boldsymbol{A}) = \sum_{j=1}^{s} \sum_{k=1}^{m_j} f_k^{(j-1)} \varphi_{jk}(\boldsymbol{A}).$$

Многочлены $\varphi_{jk}(x)$ зависят от минимального многочлена $\mu_{\boldsymbol{A}}(x)$ и не зависят от функции $f(x)$. Положим $\boldsymbol{Z}_{kj} = \varphi_{jk}(\boldsymbol{A})$. Очевидно, что эти матрицы перестановочны с матрицей $\boldsymbol{A}$.

Покажем, что матрицы $\boldsymbol{Z}_{kj}$ линейно независимы. Предположим, что $\sum_{j=1}^{s} \sum_{k=1}^{m_j} \gamma_{kj} \boldsymbol{Z}_{kj} = \boldsymbol{O}$. Рассмотрим многочлен $g(x) = \sum_{j=1}^{s} \sum_{k=1}^{m_j} \gamma_{kj} \varphi_{jk}(x)$. Тогда $g(\boldsymbol{A}) = \boldsymbol{O}$. Если не все числа γ_{kj} равны нулю, то $g(x)$ – ненулевой многочлен, так как многочлены $\varphi_{jk}(x)$ линейно независимы в силу своего определения, и $\mathsf{deg}(g(x)) < \mathsf{deg}(\mu_{\boldsymbol{A}}(x))$. Получаем противоречие с определением минимального многочлена.

Матрицы $\boldsymbol{Z}_{kj}$ называются *компонентами* матрицы $\boldsymbol{A}$. Определим компоненты матрицы в случае диагонализируемых матриц.

Теорема 41. *Пусть $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ – диагонализируемая матрица, $\mu_{\boldsymbol{A}}(x) = (x-\lambda_1)\dots(x-\lambda_s)$ – ее минимальный многочлен, $\boldsymbol{C}(x)$ – приведенная присоединенная к $x\boldsymbol{E}_n - \boldsymbol{A}$ матрица. Тогда*

$$\boldsymbol{Z}_{k1} = \frac{\boldsymbol{C}(\lambda_k)}{\mu'_{\boldsymbol{A}}(\lambda_k)}.$$

Доказательство. Очевидно, что

$$\mu'_{\boldsymbol{A}}(\lambda_k) = \prod_{j=1, j\neq k}^{s} (\lambda_k - \lambda_j) \neq 0.$$

В рассматриваемом случае

$$\varphi_{k1}(x) = \prod_{j=1, j\neq k}^{s} (x-\lambda_j) \Bigg/ \prod_{j=1, j\neq k}^{s} (\lambda_k - \lambda_j) -$$

многочлен Лагранжа. Имеем

$$\boldsymbol{Z}_{k1} = \prod_{j=1, j\neq k}^{s} (\boldsymbol{A} - \lambda_j \boldsymbol{E}_n)/\mu'_{\boldsymbol{A}}(\lambda_k).$$

Покажем, что $\boldsymbol{C}(\lambda_k) = \prod_{j=1, j\neq k}^{s}(\boldsymbol{A} - \lambda_j \boldsymbol{E}_n)$. По формуле (2) из § 5.3 $\boldsymbol{C}(x) = \Psi(x\boldsymbol{E}_n, \boldsymbol{A})$ и $(x-y)\Psi(x,y) = \mu_{\boldsymbol{A}}(x) - \mu_{\boldsymbol{A}}(y)$. Подставим вместо x число λ_k, получим $(\lambda_k - y)\Psi(\lambda_k, y) = -\mu_{\boldsymbol{A}}(y)$, так как $\mu_{\boldsymbol{A}}(\lambda_k) = 0$. Таким образом, $\Psi(\lambda_k, y) = \prod_{j=1, j\neq k}^{s}(y - \lambda_j)$ и $\boldsymbol{C}(\lambda_k) = \Psi(\lambda_k \boldsymbol{E}_n, \boldsymbol{A}) = \prod_{j=1, j\neq k}^{s}(\boldsymbol{A} - \lambda_j \boldsymbol{E}_n)$.

Теперь рассмотрим произвольную матрицу $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Имеем $\boldsymbol{A} = \boldsymbol{T}^{-1}\boldsymbol{J}\boldsymbol{T}$ для некоторой обратимой матрицы $\boldsymbol{T} \in \mathbb{C}^{n\times n}$ и жордановой нормальной формы $\boldsymbol{J}$. Далее, $\boldsymbol{J} = [\boldsymbol{J}_1, \ldots, \boldsymbol{J}_p]$, где $\boldsymbol{J}_q = \lambda_q \boldsymbol{E}_{n_q} + \boldsymbol{J}_{n_q}(0)$, $q = 1, \ldots, p$. Тогда

$$\boldsymbol{Z}_{kj} = \varphi_{kj}(\boldsymbol{A}) = \boldsymbol{T}^{-1}[\varphi_{kj}(\boldsymbol{J}_1), \ldots, \varphi_{kj}(\boldsymbol{J}_p)]\boldsymbol{T}. \tag{1}$$

Для вычисления матрицы $\varphi_{kj}(\boldsymbol{J}_q)$ рассмотрим два случая.

1. $\lambda_k = \lambda_q$. В силу определения многочленов $\varphi_{kj}(x)$ имеем $\varphi_{kj}^{(r-1)}(\lambda_k) = \delta_{rj}$ (символ Кронекера). Используя пример 4, получаем $\varphi_{kj}(\boldsymbol{J}_q) = \frac{1}{(j-1)!}\boldsymbol{J}_{n_q}(0)^{j-1}$. При этом $n_q \leq m_k$.

2. $\lambda_k \neq \lambda_q$. Тогда $\varphi_{kj}^{(r-1)}(\lambda_q) = 0$ $(r = 1, \ldots, j)$ и поэтому $\varphi_{kj}(\boldsymbol{J}_q) = \boldsymbol{O}$.

Без ограничения общности предположим, что $k = 1$. Пусть $\boldsymbol{J}_1, \ldots, \boldsymbol{J}_\ell$ – все клетки Жордана матрицы $\boldsymbol{A}$, соответствующие λ_1. Тогда из (1) получаем

$$(j-1)!\boldsymbol{Z}_{1j} = \boldsymbol{T}^{-1}[\boldsymbol{J}_{n_1}(0)^{j-1}, \ldots, \boldsymbol{J}_{n_\ell}(0)^{j-1}, \boldsymbol{O} \ldots, \boldsymbol{O}]\boldsymbol{T}. \tag{2}$$

В частности,

$$\boldsymbol{Z}_{11} = \boldsymbol{T}^{-1}[\boldsymbol{E}_{n_1+\ldots+n_\ell}, \boldsymbol{O}]\boldsymbol{T}. \tag{3}$$

Для произвольного λ_k при $k > 1$ строение матриц $\boldsymbol{Z}_{kj}$ аналогичное.

Из строения компонент непосредственно вытекает следующая

Теорема 42. *Компоненты* $\boldsymbol{Z}_{kj}$ $(k = 1, \ldots, s,\ j = 1, \ldots, m_k)$ *матрицы* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *удовлетворяют условиям:*

1) $\sum_{k=1}^{s} \boldsymbol{Z}_{k1} = \boldsymbol{E}_n$;

2) $\boldsymbol{Z}_{kp}\boldsymbol{Z}_{\ell r} = \boldsymbol{O}$, *если* $k \neq \ell$;

3) $\boldsymbol{Z}_{kj}^2 = \boldsymbol{Z}_{kj}$ *тогда и только тогда, когда* $j = 1$;

4) $\boldsymbol{Z}_{k1}\boldsymbol{Z}_{kr} = \boldsymbol{Z}_{kr}$, $r = 1, \ldots, m_k$.

Следующая теорема позволяет находить все компоненты по известным компонентам $\boldsymbol{Z}_{k1}$.

Теорема 43. *Для* $k = 1, \ldots, s$, $j = 1, \ldots, m_k$ *справедливо равенство*

$$\boldsymbol{Z}_{kj} = \frac{1}{(j-1)!}(\boldsymbol{A} - \lambda_k \boldsymbol{E}_n)^{j-1}\boldsymbol{Z}_{k1}.$$

Доказательство. При $j = 1$ доказывать нечего. Предположим, что $j \geq 2$. Без ограничения общности предположим, что $k = 1$. Положим $n' = n - \sum_{t=1}^{\ell} n_t$. Из (2) получаем

$$(j-2)!\boldsymbol{T}\boldsymbol{Z}_{1,j-1}\boldsymbol{T}^{-1} = [\boldsymbol{J}_{n_1}(0)^{j-2}, \ldots, \boldsymbol{J}_{n_\ell}(0)^{j-2}, \boldsymbol{O}_{n'\times n'}]. \tag{4}$$

Так как $\boldsymbol{T}^{-1}(\boldsymbol{A} - \lambda_1\boldsymbol{E}_n)\boldsymbol{T} = \boldsymbol{J} - \lambda_1\boldsymbol{E}_n$, имеем

$$\begin{aligned}\boldsymbol{T}(\boldsymbol{A} - \lambda_1\boldsymbol{E}_n)\boldsymbol{T}^{-1} = [&\boldsymbol{J}_{n_1}(0), \ldots, \boldsymbol{J}_{n_\ell}(0),\\ &\boldsymbol{J}_{\ell+1} - \lambda_1\boldsymbol{E}_{n_{\ell+1}}, \ldots, \boldsymbol{J}_p - \lambda_1\boldsymbol{E}_{n_p}].\end{aligned} \tag{5}$$

Перемножая почленно (5) на (4), получаем

$$\boldsymbol{T}(\boldsymbol{A} - \lambda_1\boldsymbol{E}_n)(j-2)!\boldsymbol{Z}_{1,j-1}\boldsymbol{T}^{-1} =$$

$$= [\boldsymbol{J}_{n_1}(0)^{j-1}, \ldots, \boldsymbol{J}_{n_\ell}(0)^{j-1}, \boldsymbol{O}_{n'\times n'}] = (j-1)!\boldsymbol{T}\boldsymbol{Z}_{1j}\boldsymbol{T}^{-1}.$$

Отсюда следует $(j-2)!(\boldsymbol{A} - \lambda_1\boldsymbol{E}_n)\boldsymbol{Z}_{1,j-1} = (j-1)!\boldsymbol{Z}_{1j}$, т. е.

$$\boldsymbol{Z}_{1j} = \frac{1}{j-1}(\boldsymbol{A} - \lambda_1\boldsymbol{E}_n)\boldsymbol{Z}_{1,j-1}.$$

Итерируя это равенство, получим

$$\boldsymbol{Z}_{1j} = \frac{1}{(j-1)!}(\boldsymbol{A} - \lambda_k\boldsymbol{E}_n)^{j-1}\boldsymbol{Z}_{11},$$

откуда следует требуемое.

Заметим, что для линейного оператора, заданного матрицей $\boldsymbol{A}$ на пространстве столбцов $\mathbb{C}_n$, образ оператора, заданного компонентой $\boldsymbol{Z}_{k1}$, есть в точности корневое подпространство, соответствующее собственному значению λ_k.

Рассмотрим пример. Найти компоненты матрицы $\boldsymbol{A} = \begin{pmatrix} 6 & 2 & 2 \\ -2 & 2 & 0 \\ 0 & 0 & 2 \end{pmatrix}$.

Вычислим $\chi_{\boldsymbol{A}}(x) = \begin{vmatrix} x-6 & -2 & -2 \\ 2 & x-2 & 0 \\ 0 & 0 & x-2 \end{vmatrix} = (x-2)\times$ $\times \begin{vmatrix} x-6 & -2 \\ 2 & x-2 \end{vmatrix} = (x-2)(x^2-8x+16) = (x-2)^2(x-4)$. Так как $(x\boldsymbol{E}_3 - \boldsymbol{A})\begin{pmatrix} 1 & 2 \\ 1 & 3 \end{pmatrix} = \begin{vmatrix} x-6 & -2 \\ 2 & 0 \end{vmatrix} = 4$, имеем $D_2 = 1$ и $\mu_{\boldsymbol{A}}(x) = \chi_{\boldsymbol{A}}(x)$.

Таким образом, для любой функции $f(x)$, определенной на спектре матрицы $\boldsymbol{A}$, имеем $f(\boldsymbol{A}) = f(4)\boldsymbol{Z}_{11} + f'(4)\boldsymbol{Z}_{12} + f(2)\boldsymbol{Z}_{21}$. Возьмем

$f(x) = 1$: $\boldsymbol{E}_3 = \boldsymbol{Z}_{11} + \boldsymbol{Z}_{21}$;
$f(x) = x-4$: $\boldsymbol{A} - 4\boldsymbol{E}_3 = \boldsymbol{Z}_{12} - 2\boldsymbol{Z}_{21}$;
$f(x) = (x-4)^2$: $(\boldsymbol{A} - 4\boldsymbol{E}_3)^2 = 4\boldsymbol{Z}_{21}$.

Отсюда $4\boldsymbol{Z}_{21} = \begin{pmatrix} 2 & 2 & 2 \\ -2 & -2 & 0 \\ 0 & 0 & -2 \end{pmatrix}^2 = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -4 \\ 0 & 0 & 4 \end{pmatrix}$ и

$$\boldsymbol{Z}_{21} = \begin{pmatrix} 0 & 0 & 0 \\ 0 & 0 & -1 \\ 0 & 0 & 1 \end{pmatrix}.$$

Далее, $\boldsymbol{Z}_{11} = \boldsymbol{E}_3 - \boldsymbol{Z}_{21} = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 1 \\ 0 & 0 & 0 \end{pmatrix}$ и

$$\boldsymbol{Z}_{12} = \boldsymbol{A} - 4\boldsymbol{E}_3 + 2\boldsymbol{Z}_{21} = \begin{pmatrix} 2 & 2 & 2 \\ -2 & -2 & -2 \\ 0 & 0 & 0 \end{pmatrix}.$$

§ 7.3. Последовательности и ряды матриц

Пусть $\{\boldsymbol{A}_p = (a_{ij}^{(p)})\}_1^\infty$ – последовательность матриц $\boldsymbol{A}_p \in \mathbb{C}^{n\times n}$. Говорят, что последовательность матриц $\{\boldsymbol{A}_p\}_1^\infty$ *сходится* к матрице $\boldsymbol{A} = (a_{ij})_{n\times n}$, если для всех $1 \le i, j \le n$ $\lim_{p\to\infty}(a_{ij}^{(p)}) = a_{ij}$, т. е. сходятся последовательности элементов на каждом месте матрицы.

Обозначение: $\lim_{p\to\infty} \boldsymbol{A}_p = \boldsymbol{A}$. Таким образом определена поэлементная сходимость последовательности матриц.

Очевидно, что если последовательности $\{\boldsymbol{A}_p\}_1^\infty$ и $\{\boldsymbol{B}_p\}_1^\infty$ матриц из $\mathbb{C}^{n\times n}$ сходятся соответственно к матрицам $\boldsymbol{A}$ и $\boldsymbol{B}$, то сходятся и последовательности $\{\boldsymbol{A}_p \pm \boldsymbol{B}_p\}_1^\infty$ и $\{\boldsymbol{A}_p \cdot \boldsymbol{B}_p\}_1^\infty$ соответственно к матрицам $\boldsymbol{A} \pm \boldsymbol{B}$ и $\boldsymbol{A} \cdot \boldsymbol{B}$.

Рассмотрим последовательность функций $\{f_p(x)\}_1^\infty$, каждая из которых определена на спектре матрицы $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Тогда

$$f_p(\boldsymbol{A}) = \sum_{k=1}^{s}\sum_{j=1}^{m_j} f_p^{(j-1)}(\lambda_k)\boldsymbol{Z}_{kj}, \tag{1}$$

где $\boldsymbol{Z}_{kj}$ – компоненты матрицы $\boldsymbol{A}$.

Теорема 44. *Пусть каждая из функций* $\{f_p(x)\}_1^\infty$ *определена на спектре матрицы* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *и* $\boldsymbol{A}_p = f_p(\boldsymbol{A})$, $p = 1, 2, \ldots$. *Тогда последовательность* $\{\boldsymbol{A}_p\}_1^\infty$ *сходится в том и только том случае, если каждая из скалярных последовательностей* $\{f_p^{(j-1)}(\lambda_k)\}_1^\infty$ *сходится. Если для всех* $k = 1, \ldots, s$, $j = 1, \ldots, m_k$ *и для некоторой функции* $f(x)$ *имеет место* $\lim_{p\to\infty} f_p^{(j-1)}(\lambda_k) = f^{(j-1)}(\lambda_k)$, *то* $\lim_{p\to\infty} \boldsymbol{A}_p = f(\boldsymbol{A})$. *Обратно, если предел* $\lim_{p\to\infty} \boldsymbol{A}_p$ *существует, то существует такая функция* $f(x)$, *определенная на спектре матрицы* $\boldsymbol{A}$, *что* $\lim_{p\to\infty} \boldsymbol{A}_p = f(\boldsymbol{A})$.

Доказательство. Если существуют пределы $\lim_{p\to\infty} f_p^{(j-1)}(\lambda_k)$, то согласно (1) имеем

$$\lim_{p\to\infty} \boldsymbol{A}_p = \lim_{p\to\infty} f_p(\boldsymbol{A}) = \lim_{p\to\infty} \sum_{k=1}^{s}\sum_{j=1}^{m_j} f_p^{(j-1)}(\lambda_k)\boldsymbol{Z}_{kj} =$$

$$= \sum_{k=1}^{s}\sum_{j=1}^{m_j} (\lim_{p\to\infty} f_p^{(j-1)}(\lambda_k))\boldsymbol{Z}_{kj} = \sum_{k=1}^{s}\sum_{j=1}^{m_j} f^{(j-1)}(\lambda_k)\boldsymbol{Z}_{kj} = f(\boldsymbol{A}).$$

Обратно, пусть $\lim_{p\to\infty} \boldsymbol{A}_p$ существует и $\boldsymbol{A}_p = f_p(\boldsymbol{A}) = (a_{\ell q})_{n\times n}^{(p)}$. Рассмотрим равенства (1) как систему линейных уравнений относительно m неизвестных $f_p^{(j-1)}(\lambda_k)$. Матрицы $\boldsymbol{Z}_{kj}$ определяются через матрицу $\boldsymbol{A}$ и они линейно независимы. Поэтому матрица рассматриваемой системы линейных уравнений имеет размеры $n^2 \times m$ и ранг m. Из этой системы $f_p^{(j-1)}(\lambda_k)$ определяются однозначно:

$f_p^{(j-1)}(\lambda_k) = \sum_{\ell,q=1}^n c_{\ell q}^{(k,j)} a_{\ell q}^{(p)}$, где коэффициенты $c_{\ell q}^{(k,j)}$ не зависят от p. Так как каждая последовательность $\{a_{\ell q}^{(p)}\}_1^\infty$ сходится, имеем $\lim_{p\to\infty} f_p^{(j-1)}(\lambda_k) = \sum_{\ell,q=1}^n c_{\ell q}^{(k,j)} \lim_{p\to\infty} a_{\ell q}^{(p)}$, т. е. пределы существуют. Значит,

$$\lim_{p\to\infty} \boldsymbol{A}_p = \lim_{p\to\infty} f_p(\boldsymbol{A}) = \sum_{k=1}^{s} \sum_{j=1}^{m_k} \lim_{p\to\infty} f_p^{(j-1)}(\lambda_k) \boldsymbol{Z}_{kj}.$$

Тогда в качестве функции $f(x)$ можно выбрать интерполяционный многочлен Эрмита, значениями которого на спектре матрицы $\boldsymbol{A}$ будут числа $\lim_{p\to\infty} f_p^{(j-1)}(\lambda_k)$, $k = 1, \ldots, s$, $j = 1, \ldots, m_k$.

Теорема доказана.

Пусть $\{\boldsymbol{A}_p\}_1^\infty$ – последовательность матриц $\boldsymbol{A}_p \in \mathbb{C}^{n\times n}$. Ряд матриц $\sum_{p=1}^\infty \boldsymbol{A}_p$ называется *сходящимся* к матрице $\boldsymbol{A}$, если существует предел частичных сумм $\lim_{q\to\infty} \sum_{p=1}^q \boldsymbol{A}_p = \boldsymbol{A}$. Из теоремы 44 сразу получается следующая

Теорема 45. *Если каждая функция последовательности* $\{u_p(x)\}_1^\infty$ *определена на спектре матрицы* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$, *то ряд* $\sum_{p=1}^\infty u_p(\boldsymbol{A})$ *сходится тогда и только тогда, когда сходятся ряды* $\sum_{p=1}^\infty u_p^{(j-1)}(\lambda_k)$ *для всех* $k = 1, \ldots, s$ *и* $j = 1, \ldots, m_k$. *Если ряды* $\sum_{p=1}^\infty u_p^{j-1}(\lambda_k)$ *сходятся к значениям* $f^{j-1}(\lambda_k)$ *некоторой функции* $f(x)$ *на спектре матрицы* $\boldsymbol{A}$, *то* $\sum_{p=1}^\infty u_p(\boldsymbol{A}) = f(\boldsymbol{A})$.

Следующее утверждение показывает общность данного в начале этой главы определения функции от матрицы.

Теорема 46. *Пусть матрица* $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ *имеет собственные значения* $\lambda_1, \ldots, \lambda_n$. *Пусть функция* $f(x)$ *разлагается в точке* λ_0 *в ряд Тейлора* $f(x) = \sum_{p=0}^\infty \alpha_p (x - \lambda_0)^p$, *радиус сходимости которого равен* $r > 0$. *Если* $|\lambda_j - \lambda_0| < r$ *для всех* $j = 1, \ldots, n$, *то значение* $f(\boldsymbol{A})$ *определено и* $f(\boldsymbol{A}) = \sum_{p=0}^\infty \alpha_p (\boldsymbol{A} - \lambda_0 \boldsymbol{E}_n)^p$.

Доказательство. В сделанных предположениях относительно функции $f(x)$ она имеет производную любого порядка в каждой точке λ из круга сходимости $|\lambda - \lambda_0| < r$, и эти производные получаются почленным дифференцированием ряда Тейлора для функции $f(x)$. Следовательно, $f(x)$ определена на спектре матрицы $\boldsymbol{A}$.

Положим $u_p(x) = \alpha_p(x-\lambda_0)^p$. При $p = 0, 1, \ldots$ эти функции определены на спектре матрицы $\boldsymbol{A}$ и все ряды $\sum_{p=0}^{\infty} u_p^{(j-1)}(\lambda_k)$ сходятся для всех $k = 1, \ldots, s$ и $j = 1, \ldots, m_k$. Из теоремы 45 следует, что ряд $\sum_{p=0}^{\infty} u_p(\boldsymbol{A})$ сходится к матрице $f(\boldsymbol{A})$, что и требуется доказать.

Если ряд Тейлора для функции $f(x)$ в начале координат сходится в каждой точке комплексной плоскости, то эта функция называется целой и представление для матрицы $f(\boldsymbol{A})$ в виде ряда будет сходиться для всех матриц $\boldsymbol{A} \in \mathbb{C}^{n\times n}$. Используя ряды Тейлора для тригонометрических, гиперболических и показательной функций, получаем следующие разложения, определяющие элементарные функции от матриц.

$$\sin \boldsymbol{A} = \sum_{p=0}^{\infty} \frac{(-1)^p}{(2p+1)!}\boldsymbol{A}^{2p+1}; \cos \boldsymbol{A} = \sum_{p=0}^{\infty} \frac{(-1)^p}{(2p)!}\boldsymbol{A}^{2p};$$

$$\operatorname{sh} \boldsymbol{A} = \sum_{p=0}^{\infty} \frac{1}{(2p+1)!}\boldsymbol{A}^{2p+1}; \operatorname{ch} \boldsymbol{A} = \sum_{p=0}^{\infty} \frac{1}{(2p)!}\boldsymbol{A}^{2p};$$

$$e^{\boldsymbol{A}} = \sum_{p=0}^{\infty} \frac{1}{p!}\boldsymbol{A}^{p}.$$

Если модули всех собственных чисел матрицы $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ меньше 1, то

$$(\boldsymbol{E}_n - \boldsymbol{A})^{-1} = \sum_{p=0}^{\infty} \boldsymbol{A}^p;\ \ln(\boldsymbol{E}_n - \boldsymbol{A}) = \sum_{p=0}^{\infty} \frac{(-1)^{p-1}}{p}\boldsymbol{A}^p.$$

Получив эти разложения, естественно поставить вопрос: будут ли обычные тождества между соответствующими функциями справедливы для матриц? Ответ дает следующая теорема.

Теорема 47. *Пусть $G(u_1, \ldots, u_\ell)$ – многочлен от функций $u_1, \ldots, u_\ell$ и $f_1, \ldots, f_\ell$ – определенные на спектре матрицы $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ функции, для которых функция $g = G(f_1, \ldots, f_\ell)$ равна нулю на спектре матрицы $\boldsymbol{A}$. Тогда $G(f_1(\boldsymbol{A}), \ldots, f_\ell(\boldsymbol{A})) = \boldsymbol{O}$.*

Доказательство. Обозначим через r_q интерполяционный многочлен для функции f_q на спектре матрицы $\boldsymbol{A}$ ($q = 1, \ldots, \ell$). Положим $h = G(r_1, \ldots, r_\ell)$. Так как $r_q(\boldsymbol{A}) = f_q(\boldsymbol{A})$ для $q = 1, \ldots, \ell$, имеем $G(f_1(\boldsymbol{A}), \ldots, f_\ell(\boldsymbol{A})) = G(r_1(\boldsymbol{A}), \ldots, r_\ell(\boldsymbol{A})) = h(\boldsymbol{A})$. Поскольку g и h принимают одинаковые значения на спектре матрицы $\boldsymbol{A}$, а функция g равна 0, заключаем, что многочлен h принимает нулевые значения на спектре матрицы $\boldsymbol{A}$, откуда $h(\boldsymbol{A}) = \boldsymbol{O}$, что и требуется доказать.

В частности, для любой матрицы $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ имеют место тождества

$$\sin^2 \boldsymbol{A} + \cos^2 \boldsymbol{A} = \boldsymbol{E}_n;\ \operatorname{ch}^2 \boldsymbol{A} - \operatorname{sh}^2 \boldsymbol{A} = \boldsymbol{E}_n;\ \left(e^{\boldsymbol{A}}\right)^{-1} = e^{-\boldsymbol{A}}.$$

Последнее тождество следует из того, что равенство $e^x e^{-x} = 1$ влечет за собой $e^{\boldsymbol{A}} e^{-\boldsymbol{A}} = \boldsymbol{E}_n$.

Задачи для самостоятельного решения

1. Найти компоненты матрицы:

а) $\boldsymbol{A} = \begin{pmatrix} -1 & 1 & 2 & 1 \\ 0 & -1 & 0 & -3 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \end{pmatrix}$;

б) $\boldsymbol{A} = \begin{pmatrix} 3 & 4 & 8 & 11 \\ 2 & 4 & 7 & 8 \\ -2 & -3 & -6 & -6 \\ 0 & 0 & 0 & -1 \end{pmatrix}$.

2. Вычислить значения функций $e^{\boldsymbol{A}}$, $\operatorname{sh} \boldsymbol{A}$, $\operatorname{ch} \boldsymbol{A}$ от матрицы $\boldsymbol{A}$ из задания 1.

Учебное издание

Овсянников Александр Яковлевич

ДОПОЛНИТЕЛЬНЫЕ ГЛАВЫ ТЕОРИИ МАТРИЦ

Учебное пособие

Заведующий редакцией	М. А. Овечкина
Редактор	Н. В. Чапаева
Корректор	Н. В. Чапаева
Компьютерная верстка	А. Я. Овсянников

Подписано в печать 12.02.2020 г. Формат $60\times84^{1}/_{16}$.
Бумага офсетная. Цифровая печать. Усл. печ. л. 6,28.
Уч.-изд. л. 5,5. Тираж 30 экз. Заказ 12

Издательство Уральского университета
Редакционно-издательский отдел ИПЦ УрФУ
620083, Екатеринбург, ул. Тургенева, 4
Тел.: +7 (343) 389-94-79, 350-43-28
E-mail: rio.marina.ovechkina@mail.ru

Отпечатано в Издательско-полиграфическом центре УрФУ
620083, Екатеринбург, ул. Тургенева, 4
Тел.: +7 (343) 358-93-06, 350-58-20, 350-90-13
Факс: +7 (343) 358-93-06
http://print.urfu.ru

Для заметок

ЛИДЕР УНИВЕРСИТЕТСКОГО КНИГОИЗДАНИЯ

УНИВЕРСИТЕТСКАЯ КНИГА

2018

9 785799 629472